MOTION IMAGERY STANDARDS PROFILE

Motion Imagery Standards Board

MISP-2021.1

October 2020

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHANGE LOG

2021.1	• Reviewed and Approved the following documents: MISP-2021.1, MISP-2021.1 Motion Imagery Handbook , MISP-2021.1: U.S. Governance, MISP-2021.1: U.S. Specific, ST 1910.1, ST 1909.1, ST 1801.1, ST 0605.10, ST 0601.17 • Moved prior Appendix E: Terms and Definitions to new section 1.5; Appendix F renamed Appendix E; Appendix G renamed Appendix F • Revised use of formatted text in entire document per guidance in section 1.5 • Updated Class 0 Motion Imagery to include MXF as a container • Revised Class 1 Motion Imagery to be non-specific to container type • Updated Class 2 Motion Imagery to include the MXF container • Updated Class 1/Class 2 Motion Imagery to include the CMAF container • Revised Figure 12 to add space between words • Renamed Section 6 to “Application Specific Systems” • Moved ASPA to Section 6 • Added Reqs -121 through -127 • Deprecated RP 0701 • Updated industry references and those for newly approved MISB documents

FORWARD

The Motion Imagery Standards Board (MISB) produces the Motion Imagery Standards Profile (MISP).

The MISP references technical specifications developed by the MISB and commercial industries. The MISP contains normative requirements. These requirements use the word "shall" and are mandatory for conformance to the MISP.

Commercial references cited in this document are available from industry organizations. References developed by the MISB are available under

MISB Public Web Site: http://www.gwg.nga.mil/misb or

NSG Registry Web Site: https://nsgreg.nga.mil/misb.jsp

The contents of this document are subject to continuing work within the MISB. Any changes and modifications to this document will be re-released by the MISB with an identifying change of release date when issued in a succeeding year, or an increase in version number if within the same year as follows:

MISP-YEAR.VersionNumber

NATIONAL GEOSPATIAL-INTELLIGENCE AGENCY
7500 GEOINT Drive
Springfield, Virginia 22150

July 28, 2015

Motion Imagery Standards Profile

The National Geospatial-Intelligence Agency's Office of General Counsel has determined that the standards set forth in the MISP-2015.1 and in subsequent MISB Standards and MISB Recommend Practices are not subject to the export restrictions of the Arms Export Control Act and the International Traffic in Arms Regulations.

Eric M. Fersht
Mission and International Law Division
Office of General Counsel
National Geospatial-Intelligence Agency

1 GENERAL

1.1 Scope

The Motion Imagery Standards Profile (MISP) provides requirements and general guidance of Motion Imagery Standards for the ISR community to achieve interoperability in both the communication and functional use of Motion Imagery Data. The term "shall" specify conformance requirements of this document. References cited are to specific versions documented within a given version of the MISP. The MISP states technical requirements common to the United States (U.S.) and the North Atlantic Treaty Organization (NATO) coalition partners. STANAG 4609 [1] provides NATO-specific requirements, guidance, and governance. Additional U.S.-specific requirements, guidance and governance are found in [2] and [3] respectively.

The MISP documents the Structure for data, which includes formats, encodings and containers, and the Content of data, which includes common and application-specific information that populates these structures. The Structure is based on commercial standards from Standards Development Organizations (SDO), such as ITU, ISO, SMPTE, etc., and non-commercial standards developed to support governing organizations specific activities. The content is principally based on non-commercial standards that support capability-based needs. Together the structure and content constitute Motion Imagery Data that meets conformance to the MISP.

The MISB does not issue guidance on the mechanical interconnects between sub-systems.

1.2 Objectives

The primary objective of the MISP is to further interoperability, which is key to reducing cost, effort and time when specifying, implementing, testing, and using Motion Imagery Systems.

Interoperability is the ability of systems, units, or forces to provide data, information, materiel, and services to and accept the same from other systems, units, or forces. This enables the use of data, information, material, and services exchanged to operate effectively together.

Standards provide the foundation for interoperability. Within the MISP, mandatory requirements are stated with the term "shall."

Standards and requirements alone, however, are not enough to guarantee interoperability. Rigorous testing procedures ensures mandatory standards and requirements meet conformance goals.

1.3 Composition

Figure 1 shows the ecosystem of standards supporting the MISP. This includes: MISB standards developed as new standards and profiles of existing standards; standards produced by commercial SDO's; standards produced by non-commercial SDO's; and support guidance (i.e., TRM and Motion Imagery Handbook) materials produced by the MISB. The MISB annually updates its standards along with all externally referenced standards. These updates consider cost and system capability implications to fielded and late-stage acquisition systems.

The MISP is a high-level document providing requirements for reference in acquisition and implementation activities. A companion document – the Motion Imagery Handbook [4] – supports material presented in the MISP providing definitions of terms used with more background and technical detail on the topics found here.

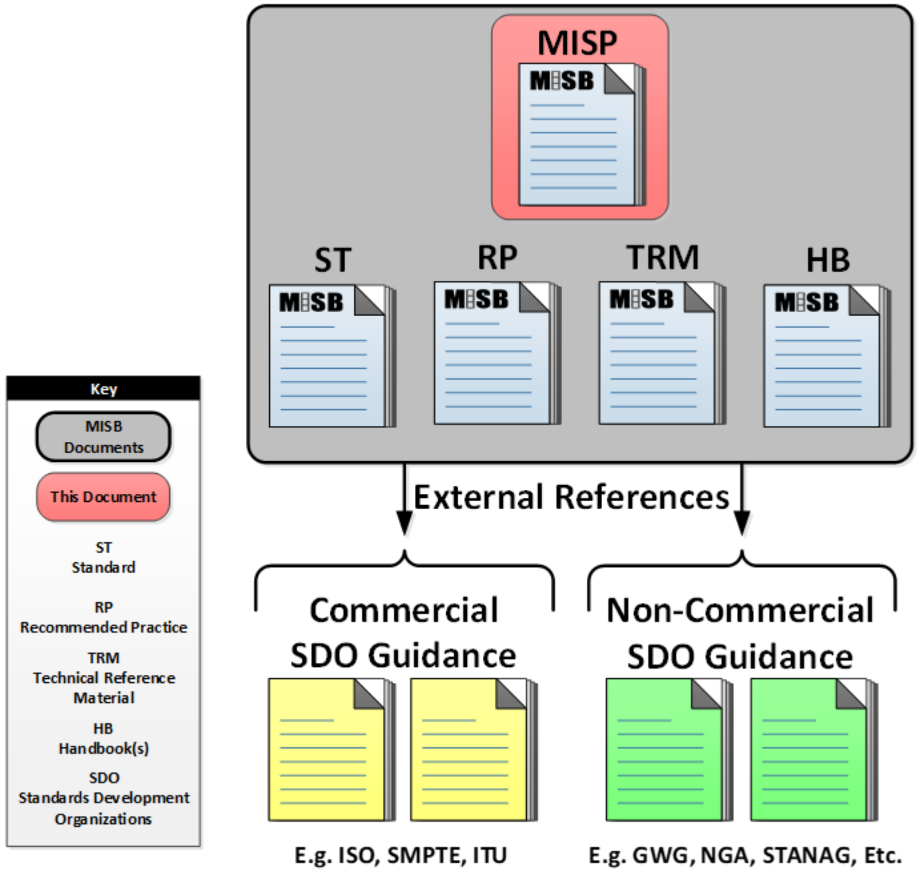

Figure 1: Ecosystem of Standards Supporting the MISP

1.4 MISB Standards Documents

The MISP leverages existing standards where possible. Often, an existing standard is too broad in scope to ensure interoperability, in which case the MISB selects a subset of its capabilities, called a profile, which becomes a MISB standard. In cases where no standard exists to meet a desired capability (i.e., a technology gap), the MISB in partnership with its community of practice develops a new one. The MISB publishes its standards documents as MISB Standards (ST) and MISB Recommended Practices (RP). See Appendix E for further information about MISB documentation.

The MISP is organized into five sections: Technical Data (Section 2), Structure (Section 3), Content (Section 4), Dissemination (Section 5) and Application Specific Systems (Section 6). The Technical Data section provides basic information and context for applying the content to the structure. The Structure section states requirements mandatory to be conformant to the MISP; these include the formatting and encoding of Motion Imagery, audio and metadata. The Content section states additional requirements on domain-specific content, such as airborne or ground-based Motion Imagery Data, that populates the structure. The Dissemination section identifies the methods and requirements used for transmission of Motion Imagery. The Application Specific Systems section provides mappings of Motion Imagery Data to external standards for inclusion into enterprise systems.

1.5 Terms and Definitions

This document defines terms and uses terms defined in the Motion Imagery Handbook. MISB terms maintain their proper noun capitalization to indicate the term is MISB defined. Overloaded terms (terms defined by more than one source which do not have a common or clear singular interpretation) will maintain their normal capitalization form (e.g., lower case, or proper noun, or acronym).

band	See Motion Imagery Handbook.
bitrate	The transmitted speed of data.
bitstream	A sequence of bits that forms the representation of coded pictures and associated data forming one or more coded Motion Imagery sequences.
chroma	Gamma-corrected chrominance signal, where gamma correction describes the total of all transfer function manipulations, such as corrections for any nonlinearities in the capture process (see SMPTE EG 28 [5] for definitions).
conformance	Confirmation by testing that a system, product, IT service, or interface adheres to a standard, standards profile, or specification.
display rate	The rate of rendering at the display in Images per second.
frame	See Motion Imagery Handbook.
image	See Motion Imagery Handbook.
imager	See Motion Imagery Handbook.
legacy system	Systems constructed from outdated technologies that may still be in use but not recommended for new development.
luma	Gamma-corrected luminance signal, where gamma correction describes the total of all transfer function manipulations, such as corrections for any nonlinearities in the capture process (see SMPTE EG 28 [5] for definitions).
MISP conformant	The ability of an implementation to create, output and/or accept Motion Imagery streams/files and recognize the component parts according to the requirements as specified the MISP.
Motion Imagery Data	Three components: Motion Imagery (see definition below), metadata and/or audio.
Motion Imagery	See MI Handbook.

Motion Imagery System	A system that provides the functionality of collecting, encoding, processing, controlling, exploiting, viewing, and/or storing Motion Imagery.
native format	Motion Imagery from an imager not subject to lossy compression.
native rate	The capture rate of an imager in Images per second of Motion Imagery.
pixel	See Motion Imagery Handbook.
protocol	System of rules for data exchange that defines syntax, semantics and synchronization of communication.
requirement	A capability required to meet an organization's roles, functions, and missions in current or future operations.
sample	See Motion Imagery Handbook.
sampling format	Specifies the horizontal, vertical, and temporal pixel density and the number of wavelength bands and bits per band for a Motion Imagery signal.
scene data	See Motion Imagery Handbook.
standard	Common, repeated, and mandatory use of rules, conditions, guidelines, or characteristics for products, or related processes and production methods; and related management systems practices.
square pixel	An image pixel where the horizontal width and vertical height are the same.
video	See Motion Imagery Handbook.

2 TECHNICAL DATA

For a Motion Imagery System to produce sharable and exploitable data, building on principles known and used by others provides a framework for interoperability. Terms, definitions, intended uses, protocols, formats, encoding, syntax and semantics must all be well defined; this includes common terms not sufficiently defined or defined differently by different communities of practice. Agreement by all stakeholders ensures successful outcomes in meeting the goals of interoperability. Standards provide an important foundation in achieving interoperability.

2.1 Motion Imagery Overview

Defining Motion Imagery and its characteristics bounds the scope of MISB and provides the framework for MISB standards. This overview suggests a common lexicon for interpreting the requirements. A Motion Imagery Functional Model facilitates relating where the standards and requirements address the various stages within a Motion Imagery System. Defining four classes of Motion Imagery assists in differentiating Motion Imagery characteristics, formats, and their respective class-specific requirements.

2.1.1 Motion Imagery Definition

Motion Imagery is a sequence of images, that when viewed (e.g., with a media player) must have the potential for providing informational or intelligence value. This implies the images composing the Motion Imagery are: (1) generated from sensed data, and (2) related to each other both in time and in space. Visible light and infrared (IR) sensed data transforms directly to an image, while synthetic aperture radar (SAR) and light detection and ranging (LIDAR) sensed data requires a conversion to a viewable image. To satisfy the time and space relationship the capture or acquisition time of each successive image must be sequentially in order <u>and</u> the space relationship between each successive image must have some recognizable visual overlap with the previous image. Figure 2 illustrates these relationships:

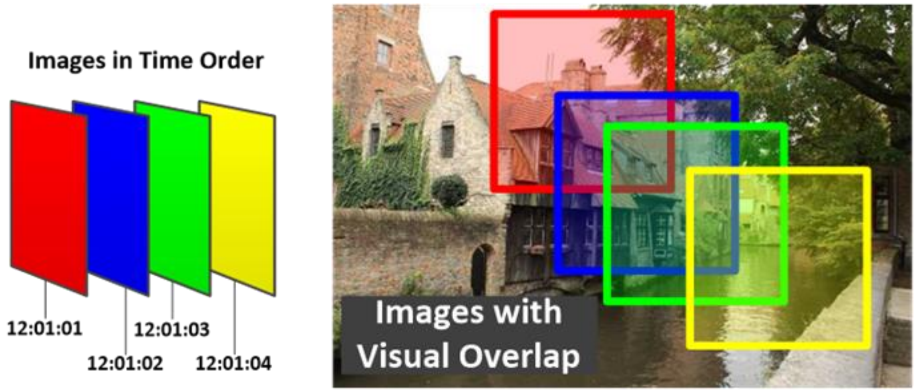

Figure 2: Illustration of Images Composing Motion Imagery
(Bruges Belgium 2014 Lat: N51 12'18" Lon: E3 13'48")

The term video implies a sequence of pictures played over time; however, unlike Motion Imagery, video does not have the same data source, time, and space restrictions. For example, video can be a simple playback of unrelated pictures, special effects, or animations that convey little if any intelligence value. Motion Imagery, on the other hand, is a restricted, highly structured form of video intended to provide information of value.

Full Motion Video (FMV) is a term used within the military and intelligence communities. As used, FMV implies a very narrow subset of Motion Imagery; one that assumes geo-spatial metadata, commercial image formats and playback rates. FMV has no formal definition and conveys different meanings to different communities; as such, the term FMV should not be used in any contractual language. For more discussion on Motion Imagery, video and FMV see [4].

Motion Imagery Systems provide the functionality of collecting, encoding, processing, controlling, disseminating, exploiting, viewing, and/or storing Motion Imagery.

2.1.2 Motion Imagery Users

To users, Motion Imagery provides a visual source of information interpreted and converted into intelligence or tactical information. Motion Imagery exploitation is the process of interpreting Motion Imagery. There are many different levels of Motion Imagery exploitation depending on the needs of the end customer and time allowed for performing the exploitation. For example, one level of exploitation might provide "situational awareness," while another might perform complex analysis with many-tiered reports and highlight clips or detailed geo-registration created. The possible outcomes for exploitation are directly linked to the capabilities of the system providing the Motion Imagery. It is important to understand and define the desired exploitation goals and requirements in the development of a Motion Imagery System. Two factors directly impact the exploitability of Motion Imagery: contextual interpretability and visual interpretability.

2.1.2.1 Contextual Interpretability

The Motion Imagery context includes all the information about the imagery, such as the where, when, and manner of collection. While a priori knowledge may prove insightful when viewing the imagery alone, metadata allows a broader audience to have contextual interpretability and is therefore preferable. Some examples of contextual information include sensor position relative to the scene; capture time of the Motion Imagery; and imager type (i.e., from visible light, IR, other).

The principal component in Motion Imagery context is the image space information, which provides the relationship of the image to the world. Specifically, it is the location of all image points in world space coordinates (e.g., latitude, longitude, height or X, Y, Z) and time. A high level of contextual interpretability is possible when the positions of all world space points in an image are known. The more accurate the image space information the higher the quality of the contextual interpretability. When the Motion Imagery context is not known, the exploitation results may become less meaningful, or even meaningless, resulting in a large cost (in time, computation and/or money) for determining this information. Motion Imagery sensors and processors are capable of measuring, computing, and embedding context information along with the Motion Imagery in the form of metadata.

2.1.2.2 Visual Interpretability

Visual interpretability is the ability to recognize objects and events within Motion Imagery. Recognition is dependent on the overall Motion Imagery quality, which is a measured ability to recognize visually what the image information represents. The atmosphere, optics, sensor, compression, delivery methods and other factors can affect Motion Imagery quality on its path to exploitation.

2.1.3 Motion Imagery Functional Model

A Motion Imagery Functional Model helps provide context for the individual MISP standards. The Functional Model organization considers the logical data flow from the Scene to Exploitation as shown in Figure 3; however, the only required function is the Imager. Note that the order of the functions is not always as indicated.

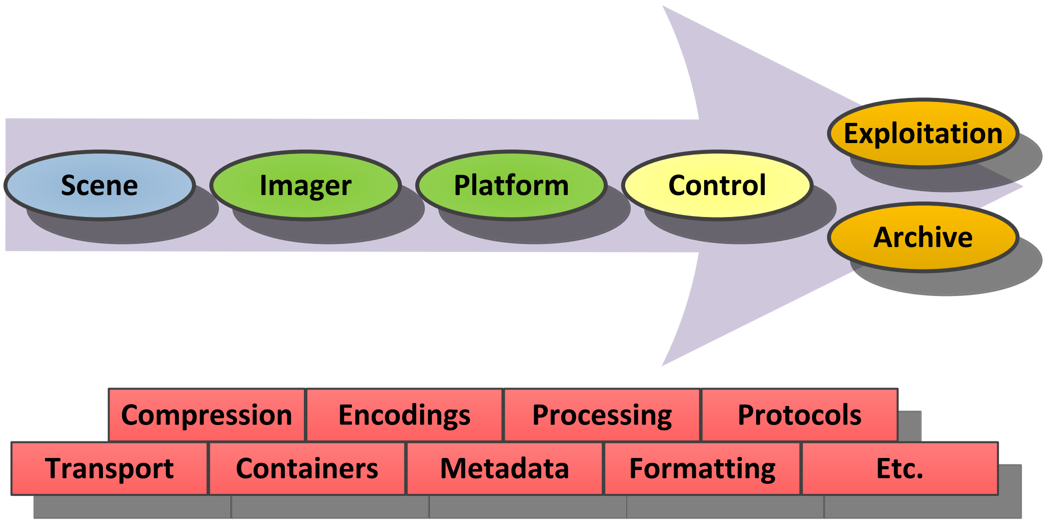

Figure 3: Motion Imagery Functional Model and Building Block Functions

This model offers a framework beginning with a Scene as imaged with an Imager through Exploitation and Archive for understanding where the standards and requirements address various elements in a Motion Imagery System. The Imager is typically a subsystem on a Platform, which could be airborne, ground, sea or other, and under user direction via a Control means. Collected Motion Imagery supplies various phases of Exploitation and Archival. Table 1 summarizes the Functional Model components with more details found in [4].

Table 1: Functional Model Components

Component	Description
Scene	The scene is the area viewed by an imager. When an imager is imaging a scene it typically measures a set of one or more data samples from one or more different modalities, such as visible light, infrared, LIDAR, etc., that radiate or reflect from the objects in the scene.
Imager	The imager converts information from a scene into an image, and when possible provides support information about the imager characteristics and time of image capture.
Platform	A platform is a static or movable system supporting an imager (sensor). A platform may provide information regarding its environment in the form of metadata.

Control	A control device directs the imager (and platform) position, orientation, or other attributes. Motion Imagery Systems generally allow for control of the imager, whether orienting its direction dynamically (i.e., pan/tilt), or modifying its parameters, such as contrast, brightness, image structure, zoom, etc.
Exploitation	Exploitation is the end use of Motion Imagery and ranges from simple situational awareness – the when and where, to in-depth extraction of detected features, measurement, and coordination with other intelligence data. Exploitation may be by human, machine, or both, and may occur in real-time or post collection.
Archive	Archiving of all Motion Imagery from the sensor and additional exploitation data for later phases of exploitation, generating reports and historical reference. An important aspect of storage is file format. A standardized file format for the search and discovery of Motion Imagery is critical to reuse.

In support of the Functional Model, there are several utility building block functions. The building block functions populate a common toolbox of standards that as a sum-of-parts provide a given capability. Requirements on these functions assure optimal interoperability when operating together and when used across various implementations. The MISP identifies requirements at the Functional Model level that facilitate interoperability of the building block functions. Table 2 lists the most common types of building block functions.

Table 2: Building Block Functions

Building Block Function	Description
Compression	Compression is an algorithmic sequence of operations to reduce the quantity of data in non-compressed Motion Imagery.
Encodings	Encodings are specific representations of data. As an example, KLV (Key Length Value) is an efficient and extensible representation of metadata.
Format	Formats specify the organization or arrangement of data.
Container	Containers encapsulate one or more data in a defined format. Example: the standardized MPEG-2 Transport Stream container can hold Motion Imagery of a supported format, and metadata of its own format.
Protocols	Protocols provide the rules for systems to communicate. For example, a protocol defines the interface specifications for data transfer between functions along the Motion Imagery Functional Model.
Processing	Processing modifies or augments the Motion Imagery. Processing includes data type conversions, sampling format changes and enhancement algorithms. Examples include image transformations, and infrared data range scaling.

2.1.4 Motion Imagery Anatomy

As defined in Section 2.1.1, Motion Imagery is a temporal sequence of images, which overlap spatially, as illustrated in Figure 4.

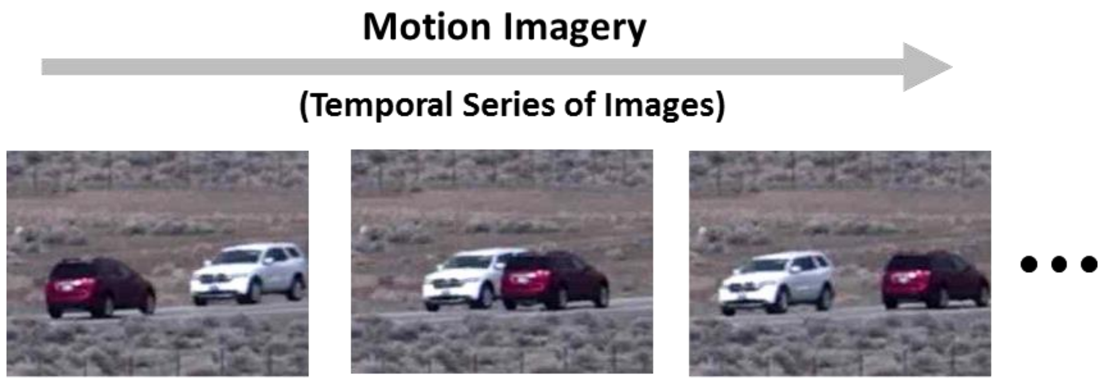

Figure 4: Sequence of Images
(Copyright (C) 2014 Her Majesty the Queen in Right of Canada as represented by the Minister of National Defence. All rights reserved.)

Each image in the sequence represents a defined time of image capture. The native rate is the capture rate of each image measured in images per unit of time (e.g., seconds, minutes, etc.). The display rate is the rendering rate of images on a display. The native rate and display rate do not necessarily have to be the same. For example, by skipping images the display rate is faster than the native rate; conversely, by repeating images the display rate is slower than the native rate.

> Note on terminology: The Motion Imagery Handbook defines the term image as a frame with pixels derived from a set of sensed phenomena. Because industry uses the term frame when quoting the number of images per second i.e., frames per second, both are equivalent in this document.

2.1.4.1 Image Characteristics

An image is a rectangular array of pixels, as shown in Figure 5. Two characteristics define this array: the number of pixels per image, which is the product of the number of columns of pixels by the number of rows of pixels, and the image aspect ratio, which is the ratio of the number of columns to the number of rows of pixels.

Each pixel in an image is a computed value using the data from one or more measurement arrays, called bands. The number of bands is the count of bands used to compute the pixels in the image and is a constant value for the whole image. For example, a grey-scaled image is a one-band image; color images are three-band images (e.g., Figure 5); multi- and hyper-spectral imagery have many bands.

Each band is an array of samples that represent measured phenomena, for example, the amount of red, green, or blue light, as illustrated in Figure 5. The number of samples in a band is equal to the number of columns times the rows of samples in that band. Each band in an image may have a different number of samples; thus, to form a pixel, multiple pixels may share sample data, such as in 4:2:2 and 4:2:0 formats (see Section 2.1.4.2). Summing the number of samples per band produces the number of samples per image.

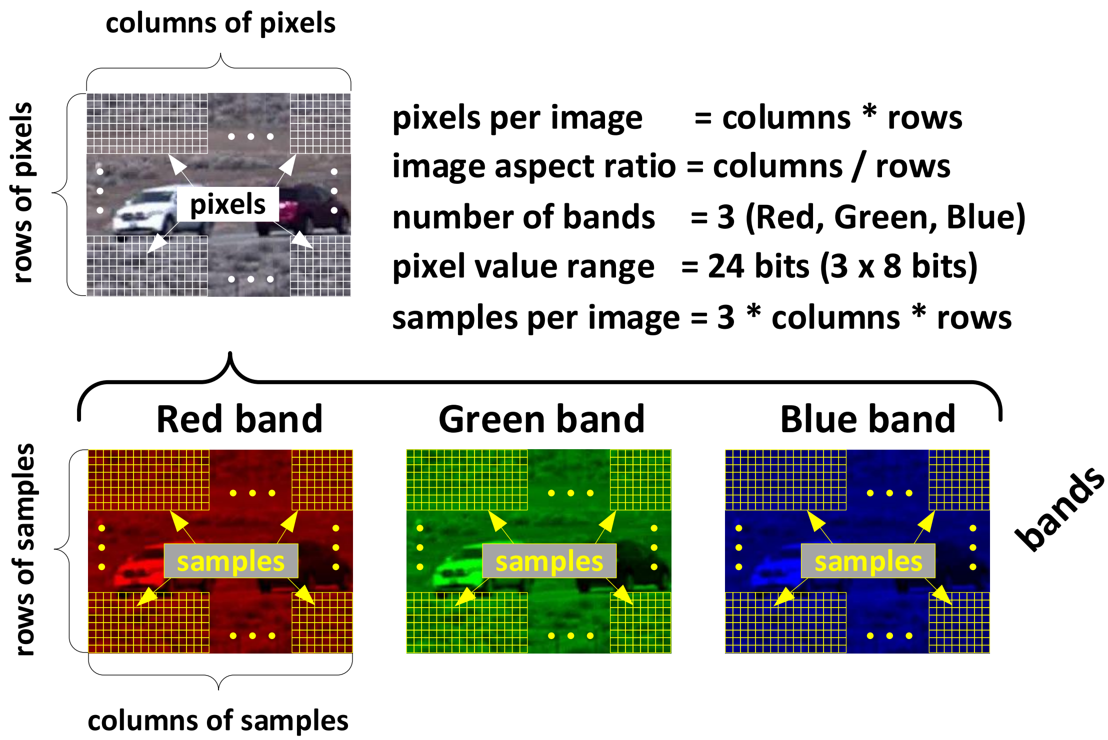

Figure 5: Anatomy of an Image Using Color as an Example
(Copyright (C) 2014 Her Majesty the Queen in Right of Canada as represented by the Minister of National Defence. All rights reserved.)

In a process called quantization, an imager produces samples within a numerical data range bounded by a minimum and maximum value and divided into intervals. The number of intervals, or quantization, is determined by the number of bits used to represent the numerical data range. For example, 8-bit quantization will have 256 (2^8) intervals, while 10-bit quantization will have 1024 (2^{10}) intervals. Dividing the data range into equal intervals is the most common quantization found in Motion Imagery and audio signal processing.

The sample value range expressed in bits per sample (bps) is a quantized value of a measured phenomenon. Each band's sample value range can be different from other bands in the same image.

Each pixel in an image represents a combination of several samples collectively taken from the number of bands in an image. The pixel value range is the total number of bits that represent the collective samples. For example, a color image with 8 bits in each of three bands would have a pixel value range of 24 bits. Likewise, an image with 10 bits per band and 5 bands would have a pixel value range of 50 bits.

Determining the pixel value range of multi-band imagery is straightforward when all samples in all bands have the same number of bits. Determining the pixel value range of three-band color imagery is based on the color sampling model.

2.1.4.2 Color Model

A color model is a mathematical representation of colors, for example, expressed with a triplet of bands. For example, RGB (Red, Green, and Blue) describes which wavelengths of light produce a given color. Color model sampling describes the samples required to represent each color band, and the notation used is three numbers separated by colons, for instance 4:4:4, 4:2:2, 4:2:0, etc. In the 4:4:4 color model, each color band contains the same number of samples (i.e., the same sampling). Three samples – one sample from each color band represent a color pixel.

The human eye, however, is more sensitive to image brightness than color. Some color models leverage this to reduce the data needed to convey color. Two common color models – 4:2:2 and 4:2:0 – transform the RGB color bands into a luma (brightness information) band and two chroma (color information) bands. The luma band is a linear combination of the RGB bands. Each chroma band is a difference of colors, and so in the absence of color a chroma band is zero. Because the human eye is more sensitive to brightness, the neighboring color-difference samples when averaged share the value with more than one brightness sample; this reduces the number of color-difference values (i.e., samples) in an image.

These transformations and the sharing reduce the quantity of samples for the three bands by one-third (for 4:2:2) and one-half (for 4:2:0). Often these transformations precede an image compression operation. The Motion Imagery Handbook provides more detail on color models.

2.1.4.3 Aspect Ratio

Image aspect ratio is the ratio of image width to height, i.e., 16:9, 4:3, etc. However, this definition does not completely characterize why some images when viewed are geometrically incorrect, or a portion of an image cut off. Depending on how the internal display processing interprets the received format different displays show different geometric effects. Understanding these effects is important in exploitation when performing mensuration or using automated analyses processes. The terms Pixel Aspect Ratio (PAR), Source Aspect Ratio (SAR) and Display Aspect Ratio (DAR) help understand these effects.

Each pixel created by an imager has a height and width dimension, usually measured in micrometers. PAR is the ratio of the width divided by the height. A square pixel has a 1:1 ratio. The user of image data must be aware of the PAR in both preprocessing, display, and post-processing. In pre-and post-processing, improper understanding of PAR can cause inaccurate data resulting in incorrect analysis and measurement. In display, improper accounting of PAR can produce visible geometrical distortion.

The SAR is the image aspect ratio acquired at the imager. This could be the imager's native format, or a converted format (see Figure 8). For example, a high-definition imager has a SAR of 16:9. The MISB specifies the SAR be in the range of [0.25, 4.0], which is effectively a 1:4 to a 4:1 ratio limit. MISB TRM 1404 [6] discusses this reasoning.

Requirement	
MISP-2015.1-01	The Motion Imagery Source Aspect Ratio (SAR) shall be in the range [0.25, 4.0] (see MISB TRM 1404).

The DAR is the ratio of horizontal to vertical dimensions of the display area.

The relation between these three metrics is SAR x PAR = DAR, and this relationship helps understanding source-to-display image irregularities. MISB RP 0904 [7] provides additional information on SAR, DAR and PAR, and guidance for scaling and cropping Motion Imagery.

2.1.4.4 Scanning Type

Scanning type defines the method an imager uses to expose, capture, and form an image. Three scanning types are common: progressive-global, progressive-rolling (also called rolling shutter) and interlaced. With progressive-global scanning, exposure, capture and formation of the entire image occurs simultaneously. In progressive-rolling scan, exposure and capture of the image is sequentially row by row forming a final image. In interlaced scan, exposure and capture of an image occurs in two passes. In a first pass, exposure and capture occur for the odd numbered rows followed by a second pass, where exposure and capture occur for the even numbered rows; interleaving the rows from each half-captured image together forms a complete image.

Figure 6 shows two images and the possible distortions that can occur to each image for each scanning type. The first image illustrates the distortions that can occur with moving objects (vehicle) within the scene. The second image illustrates the distortions that can occur when the sensor is moving. For either image, when using progressive-global scan the object shows no distortion. With progressive-rolling scan the scan is from top to bottom of an image, so the object appears tilted and distorted. With interlaced scan, because of the odd row / even row capture process, the edges of the object are torn and displaced where there is motion within the frame.

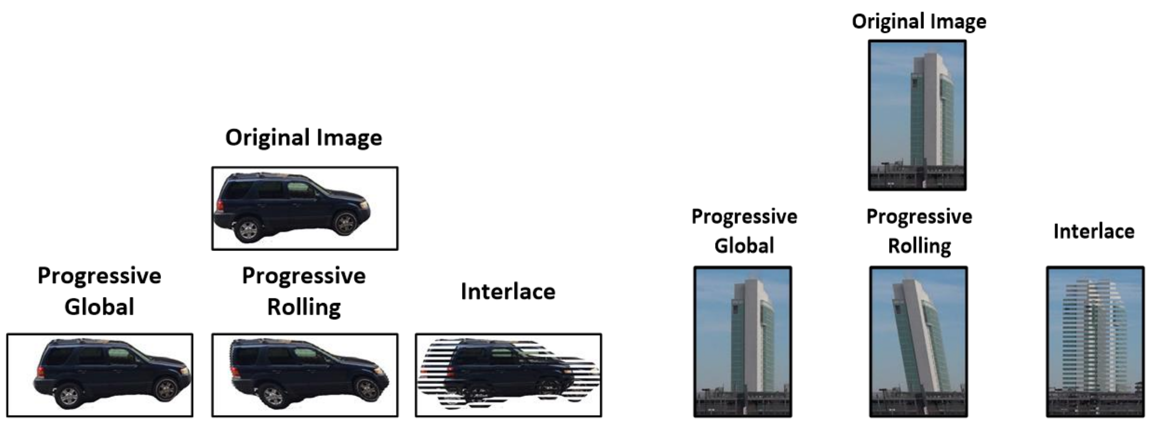

Figure 6: Examples of Progressive-Global, Progressive-Rolling and Interlaced Scan

Progressive-global scan does not introduce temporal artifacts like progressive-rolling and interlaced scan, and therefore, is far superior in faithfully capturing motion. Eliminating temporal artifacts, such as edge tearing, facilitates image exploitation and improves compression efficiency.

Given the obvious advantage of progressive-global over progressive-rolling, the MISB <u>recommends</u> the use of progressive-global scan. Given the obvious advantage of progressive scan over interlaced scan, the MISP <u>mandates</u> the use of progressive scan (either progressive-global or progressive-rolling) for all types of Motion Imagery.

Requirement	
MISP-2015.1-02	The Motion Imagery format at the imager (Sensor) and for any Motion Imagery conversions or transcodes shall be a progressive-scan format.

This requirement specifically bans the use of interlaced scan formats, such as NTSC [8], PAL [9], SECAM [9], and High Definition 1080i [10] by any MISP conformant system, this ban includes interlaced scan formats at the imager and for any Motion Imagery conversions or transcodes.

The Motion Imagery Handbook provides additional information on scanning types. MISB ST 1507 [11] provides a metadata model to represent timing information within a global and a rolling shutter imager.

2.1.4.5 Sampling Format

Many sensors use commercial camera technologies to produce Motion Imagery which adheres to specific formats. The industry term sampling format describes the Motion Imagery format. The following list of characteristics defines a sampling format:

- Number of samples in a band
- Number of bands in an image
- Number of pixels in an image
- Color model (if used)
- Number of images per second
- Scanning type
- Aspect ratio
- Vertical image size (number of rows in an image)
- Horizontal image size (number of columns per row in an image)

2.1.5 Motion Imagery and Time

Time is a critical aspect of Motion Imagery. Time forms the basis of capturing and displaying Motion Imagery based on the sampling format selected. Time links different sources of information for synchronizing the presentation of data, such as Motion Imagery, metadata and audio.

Two types of time information – Absolute Time and Relative Time – serve different purposes. Absolute Time is based on a well-defined reference source, such as International Atomic Time (TAI), and enables coordination amongst sensing systems that collect information (i.e., Motion Imagery, metadata, and other information sources.) Relative Time, based on an internal or localized timing reference which may be independent of an external reference, helps to maintain synchronizing data for rendering on a display, (i.e., displaying Motion Imagery, metadata and audio simultaneously).

Both the MISP Precision Time Stamp and the Nano Precision Time Stamp defined in MISB ST 0603 [12] represent Absolute Time. The presence of a timestamp based on Absolute Time is mandatory for all Motion Imagery; it is also mandatory for metadata packets which include a metadata item for a timestamp based on Absolute Time. Regardless of the representation of timestamp used the same representation guarantees consistency throughout an instantiation of Motion Imagery. Thus, a timestamp representation cannot change within an instantiation of Motion Imagery.

Requirement(s)	
MISP-2018.3-116	Every Motion Imagery *frame* shall include a timestamp representing Absolute Time consistent with MISB ST 0603.
MISP-2018.1-97	Where metadata contains a timestamp item representing Absolute Time, the timestamp shall be in accordance with MISB ST 0603.
MISP-2018.1-98	An instantiation of Motion Imagery shall have only one timestamp representation which represents Absolute Time.

MISB ST 0603 also defines one type of Relative Time measure called a Commercial Time Stamp – equivalent to industry standard time code. The Commercial Time Stamp is optional in MISP Motion Imagery.

2.2 Motion Imagery Classes

Motion Imagery spans a large set of imagery types and sampling formats. Defining Motion Imagery classes helps group Motion Imagery per their own unique methods for creation, delivery, and storage. Four of the characteristics described in Section 2.1.4 help to classify Motion Imagery: pixel value range, number of bands, number of pixels per image and number of images per second. Many sources produce Motion Imagery, and these sources require certain sampling formats. In consideration of this varied range of sampling formats and potential use cases, the MISP segments Motion Imagery into classes (see Figure 7) that cite class-specific requirements and standards as follows:

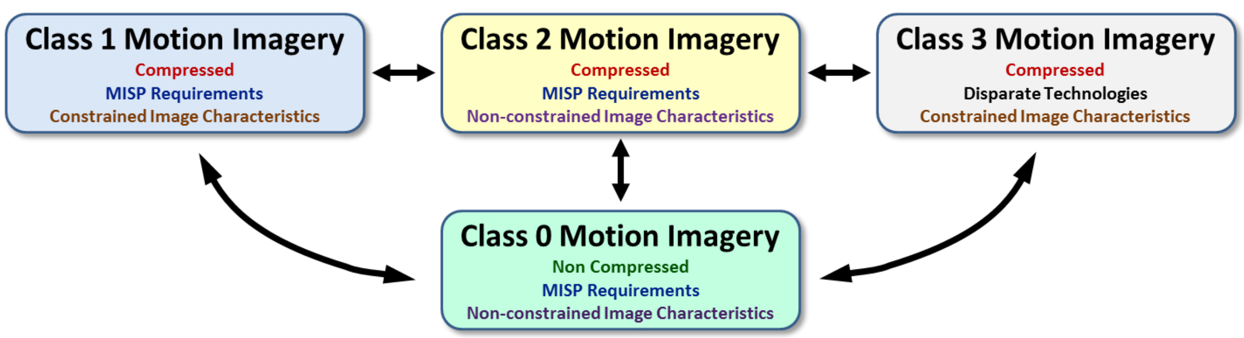

Figure 7: Classes of Motion Imagery

Class 0 Motion Imagery represents the collective requirements for non-compressed Motion Imagery, metadata, audio, and suitable containers. Class 0 Motion Imagery allows for non-constrained image characteristics (see Sec. 2.1.4.1), such as unrestricted pixel value range, number of bands, number of pixels per image and number of images per second. Examples include 14-bit infrared, 10-band Hyperspectral (HSI), and 3-band High Definition.

Class 1 Motion Imagery represents the collective requirements for compressed Motion Imagery, metadata, audio, and suitable containers. Class 1 Motion Imagery is applicable when delivering monochrome and color Motion Imagery in cases where the transmission bandwidth prohibits the use of Class 0 Motion Imagery. Typically, the image characteristics of Section 2.1.4.1 are constrained; that is, limits in pixel value range, number of bands, and number of images per second. Example: streaming of H.264/AVC (Advanced Video Coding) compressed airborne Motion Imagery. Class 1 Motion Imagery, based on standards from commercial SDO's, use three bands of color.

Class 2 Motion Imagery represents the collective requirements for compressed Motion Imagery, metadata, audio, and suitable containers. Unlike Class 1 Motion Imagery, Class 2 Motion Imagery allows for a non-constrained set of image characteristics. Examples include high frame rate scientific imaging, Large Volume Motion Imagery (LVMI), high bit-depth compressed infrared. Class 2 Motion Imagery is based on standards from commercial SDO's and non-commercial standards unique to a governing organization.

Class 3 Motion Imagery represents sources external to the NSG, such as cell phones, mobile devices, and surveillance cameras, which may use formats, compression, containers, and other technologies that do not conform to the MISP. The requirements specified for Class 3 Motion Imagery address conversion of Motion Imagery, metadata, audio, and containers to meet Class 1 Motion Imagery requirements.

The arrows in Figure 7 denote conversion amongst the classes of Motion Imagery. Non-compressed Class 0 Motion Imagery provides a common denominator for the Class 1/Class 2/Class 3 compressed forms of Motion Imagery. Conversion from one class to another may require a change in image characteristics: for example, number of image bands, bit depth, spatial density, frame rate, etc. Typically, conversion necessitates first decompressing the Motion Imagery back to a Class 0 Motion Imagery format, followed by subsequent compression per the characteristics of the class selected. The arrows represent notional conversion rather than direct conversion.

2.3 Motion Imagery Domains

Motion Imagery Domains are application areas of Motion Imagery that have their own unique characteristics; for example, Airborne, Ground, Maritime, Surface, Underwater, Combat, Visual Light, Infrared (IR) and Synthetic Aperture Radar (SAR) are all different domains. Domains provide a method of grouping common items for developing a set of Standards or Recommended Practices. Domains can overlap, for example an IR-Combat-Airborne Motion Imagery System will have characteristics of IR, Combat, and the Airborne Domains; therefore, standards that apply to all three domains are applicable.

3 STRUCTURE

Standards from SDO's, such as ITU, ISO, SMPTE, etc., and non-commercial standards developed to support governing organizations specific activities rely on defined structures to Motion Imagery. The Motion Imagery classes identify which standards apply and add requirements needed to support general application areas. Applications which span classes require guidance to convert one class to another.

Several enduring principles underlie the path forward towards improving interoperability and image quality, both which will lead to improved information gathering and exploitation.

3.1 Enduring Principles

The following is a roadmap for continued improvement of the technical aspects of Motion Imagery capabilities across the enterprise:

- Optimize Motion Imagery spatial, temporal, and spectral dimensions to meet performance objectives.
- Maximize and retain image quality throughout the entire image workflow given cost and infrastructure constraints.
- Incorporate threshold metadata to support basic situational awareness, discovery & retrieval, and cross-domain dissemination and objective metadata to support higher fidelity exploitation based on system requirements.
- Consider expected performance improvements versus life cycle costs and impacts to infrastructure when upgrading to new compression and format standards.

3.2 Common Attributes and Requirements

There are many methods for representing Motion Imagery, metadata, and audio information. Of these, most critical to Motion Imagery exploitation is the visual Motion Imagery. It is the metadata, however, that provides the essential contextual information (see Section 2.1.2.1), which greatly increases the value of the Motion Imagery. The container is the unifying package that holds the Motion Imagery and metadata. Motion Imagery, metadata, audio, and container each have a defined structure and encoding for its data. This structure forms the basis of interoperability.

3.2.1 Motion Imagery

There are two basic representations for Motion Imagery: analog and digital. Although analog representations for moving pictures underscored the beginnings of commercial television, digital representations supply great advantages in image quality, usability, and storage. The intent of the following requirements is to leverage these advantages by mandating digital representations of Motion Imagery across all classes of Motion Imagery. Legacy Motion Imagery is Motion Imagery found in systems that pre-date the MISP requirements for digital Motion Imagery.

Requirement(s)	
MISP-2015.1-05	Legacy Motion Imagery in analog form shall be converted into digital form.
MISP-2015.1-06	Motion Imagery in digital form shall remain in digital form.

Non-compressed digital Motion Imagery produces a large quantity of data. Compression technology reduces this data to enable the usability of Motion Imagery across applications. Motion Imagery compression is an algorithmic process designed to eliminate redundant data by trading the degree of compression for image quality. There are two types of compression: temporal and spatial. Temporal compression removes redundancies across a time sequence of images. Spatial compression removes localized redundancies within an image itself. An encoder executes a compression algorithm and produces a compressed signal in a specified format called a bit-stream. Encoders may perform spatial compression only, or a combination of both spatial and temporal compression.

Decompression reverses this process by converting compressed Motion Imagery back to a non-compressed state. Because compression generally reduces image quality, the steps of compression and decompression are not perfectly reversible; that is, the decompressed image will have lower image quality than the image prior to compression. While there are visually lossless compression technologies, the degree of data reduction possible is not enough to meet many common user requirements. A decoder performs decompression, in this case, decompressing Motion Imagery.

3.2.2 Metadata

KLV (Key-Length-Value) is a very efficient method for encoding metadata. The following requirements apply to all metadata within the MISP developed as KLV. SMPTE ST 336 [13] and MISB ST 0107 [14] govern the encoding and use of metadata. MISB ST 1201 [15] provides a mapping between floating point and integer values. MISB ST 1303 [16] defines how to express a multi-dimensional array in KLV.

Requirement(s)	
MISP-2015.1-07	KLV (Key-Length-Value) metadata shall be encoded in accordance with SMPTE ST 336.
MISP-2015.1-08	KLV metadata shall be formatted in accordance with MISB ST 0107.
MISP-2015.1-09	When using KLV and mapping between floating point values and integer values, the mapping shall comply with MISB ST 1201.
MISP-2015.1-10	Multi-dimensional arrays of data expressed in KLV shall be formatted in accordance with MISB ST 1303.

3.2.3 Container

A container plays an important role in furthering interoperability. Too many container types place undue burden on systems in unwrapping and rewrapping data. The “optimal” container serves a wide range of capabilities and provides growth for new capabilities. The following requirement forces the use of MISP-approved containers.

Requirement	
MISP-2015.1-11	Only those containers specified in this document (MISP) shall be used for Class 0 Motion Imagery, Class 1 Motion Imagery, and Class 2 Motion Imagery.

3.3 Single Image Extraction

An image extracted from Motion Imagery is no longer Motion Imagery. Such extractions could be in JPEG, BMP, JPEG2000 or other image formats, or directly as a NITF (National Imagery Transmission Format) [17]/NSIF (NATO Secondary Image Format) [18]. An image extracted from Motion Imagery into one of these formats is by direct conversion and storage using no transitional analog processing steps.

Requirement	
MISP-2015.1-16	Digital images extracted from Motion Imagery as a NITF (National Imagery Transmission Format)/NSIF (NATO Secondary Imagery Format) shall comply with MIL-STD-2500 / STANAG 4545 respectively.

3.4 Class Conversion

Motion Imagery captured by an imager has a defined native format that falls into one of the classes of Motion Imagery. In typical applications, use of more than one class is likely. For instance, processing an imager producing non-compressed High-Definition Motion Imagery into Class 1 Motion Imagery for dissemination.

Figure 8 illustrates common conversions for converting Class 0 Motion Imagery to a different class as needed. For example, the "Embedded" process may convert one or more instances of Class 1 Motion Imagery into a Class 2 Motion Imagery data set, such as an array of cameras producing Class 1 Motion Imagery stored in a Class 2 Motion Imagery file format. The "Window" process might represent an extracted Class 1 Motion Imagery region-of-interest from Class 2 Motion Imagery, perhaps according to a High-Definition format. MISB RP 1011 [19] provides guidance for H.264/AVC compression and dissemination from a Class 2 Motion Imagery platform.

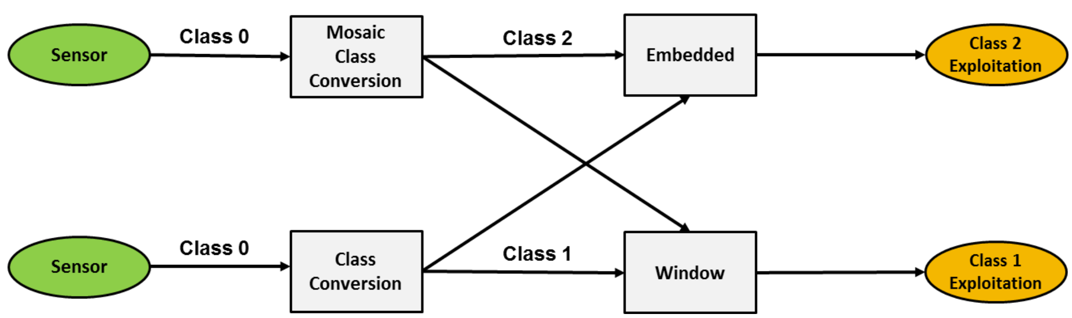

Figure 8: Class 0 Motion Imagery Conversion

When converting Motion Imagery, the imagery may have its characteristics changed; for example, converting Class 2 Motion Imagery with a high pixel value range to Class 1 Motion Imagery via pixel value range scaling, or Class 2 Motion Imagery with a high number of images per second to Class 1 Motion Imagery via dropping images.

3.5 Class 0 Motion Imagery

Class 0 Motion Imagery describes Motion Imagery in its non-compressed form as output by an imager. Non-compressed Motion Imagery has advantages over compressed Motion Imagery:

higher image fidelity with fewer artifacts, which affects feature extraction and other image processing algorithms; fewer prolonged artifacts when an error in the data occurs; and accurate edits when authoring clips. One disadvantage of non-compressed Motion Imagery is the cost in managing, moving, and storing such a large quantity of data.

Processes like noise reduction, stabilization, etc. improve the usability of non-compressed Motion Imagery. Some processes, such as pixel value range scaling used in compressing infrared Motion Imagery, reduce the detail in the Motion Imagery. Moving computationally intensive operations as these prior to compression should be an objective as the SWaP (Size/Weight and Power) of processing systems at the source allow.

3.5.1 Sampling Format

Many SDO's publish standards that define the sampling format and interface to an imager. The following sections cite commercial industry standards and state requirements applicable to the sampling format of Class 0 Motion Imagery.

Table 3 lists sampling formats for Class 0 Motion Imagery with their respective governing industry standard.

Table 3: Sampling Formats of Referenced Commercial Standards

Sampling Format	Native Rate (Hz)	Standard	Aspect Ratio (SAR)	Color Sampling	Quantization Bits/Sample
525p[1]	60M	ITU-R BT.1358 [20]	4:3, 16:9	4:2:2, 4:4:4	8 or 10 (uniform)
625p	50				
720p	60	SMPTE ST 296 [21]	16:9	4:2:2, 4:4:4	8 or 10 (uniform)
	60M				
	50				
	30				
	30M				
	25				
1080p	60	SMPTE ST 274 [10]			8, 10 or 12 (uniform)
	60M				
	50				
	30				
	30M				
2160p	120, 60, 60M, 50, 30, 30M, 25	SMPTE ST 2036-1 [22]	16:9	4:2:2, 4:4:4	10 or 12 (uniform)
4320p					

[1]Note: "p" denotes progressive scan
[2]M = 1000/1001

The Sampling Format column lists the common industry designation for the format. The Native Rate column is the number of images per second. The Standard column indicates the governing standard for sampling format. The Aspect Ratio column is the SAR, see Section 2.1.4.3. The

Color Sampling column references the applicable color sampling model, see Section 2.1.4.2. The Quantization column is the sample value range, see Section 2.1.4.1.

3.5.1.1 Sample Value Range

The sample value range represents the number of bits per sample to represent faithfully scene information and is dependent on both the range (i.e., lowest to highest value) of the information and the desired resolution (i.e., number of discrete values). For example, the sample value range of a visible light image is typically 8-bits. The sample value range of an IR image is greater; up to 14 bits. The sample value range becomes a factor in subsequent stages of processing, such as compression where equipment that can only accept 8-bit data is common. IR data greater than 8 bits requires scaling to fewer bits per sample. Scaling the sample value range can compromise the integrity of the data. MISB ST 0404 [23] discusses methods for scaling IR data.

3.5.1.2 Pixel Value Range

Adding the sample value range for all bands at a pixel position gives the pixel value range for Class 0 Motion Imagery. For a monochrome image, such as single band IR with a sample value range of 14 bits, the pixel value range is 14 bits. For a three-band RGB image with a sample value range of 8 bits per band, the pixel value range is 24 bits (3 bands x 8 bits/band).

When RGB color is transformed into a luma and chroma model, the average pixel value range reflects the reduced number of samples in the chroma bands; as discussed in Section 2.1.4.2, where the average pixel value range for 4:2:2 is 16 bits, and for 4:2:0 is 12 bits.

3.5.1.3 Display Aspect Ratio and Pixel Aspect Ratio

Class 0 Motion Imagery includes the common industry-defined sampling formats of Enhanced Definition (ED), High Definition (HD) and Ultra-High Definition (UHD). The ED format is defined in a 4:3 display aspect ratio (i.e., DAR = 4:3), where the image has a horizontal to vertical ratio of 4:3. HD and UHD formats specify a 16:9 display aspect ratio (i.e., DAR = 16:9).

In general, ED formats may or may not have square pixels depending on the image sampling format. Square pixels (i.e., PAR = 1:1) are a characteristic of the HD and UHD standards consistent with square-pixel computer displays.

3.5.1.4 Native Data Rate

Appreciating the quantity of data produced by an imager is an important metric when estimating the degree of compression needed to meet system constraints. The native data rate of Motion Imagery for a given sampling format is:

pixels per image = (number of horizontal pixels) x (number of vertical pixels)

native image size (bytes) = (pixels per image) x (pixel value range) / (8 bits per byte)

native data rate (bytes/sec) = (native image size) x (number of images per second)

Table 4 lists example native data rates. Column Pixels and Row Pixels indicate the image dimensions. Pixels per Image is the product of the image dimensions. The Native Image Size is the product of the Pixels per Image and the Pixel Value Range. Finally, Native Data Rate is the product of the Native Image Size and the Number of Images per second. Note that the overall data rate must consider the overhead from metadata, audio, and the container.

Table 4: Example Native Data Rates

Column Pixels	Row Pixels	Pixels Per Image	Pixel Value Range (bits)	Native Image Size (MB)	Image Rate (Images/sec)	Native Data Rate (MB/sec)
1280	720	921,600	24 (4:4:4)	2.765	60	165.9
1280	720	921,600	16 (4:2:2)	1.843	60	110.6
1280	720	921,600	12 (4:2:0)	1.382	60	82.9
1920	1080	2,073,600	12 (4:2:0)	3.110	30	93.3

3.5.1.5 Commercial Sampling Formats

The following requirements apply when using the sampling formats for ED, HD and UHD. The standards referenced are those from commercial SDO's.

3.5.1.5.1 Enhanced Definition

Requirement	
MISP-2015.1-19	The Motion Imagery sampling format for Enhanced Definition (ED) Motion Imagery shall be defined by ITU-R BT.1358.

3.5.1.5.2 High Definition

Requirement(s)	
MISP-2015.1-20	The Motion Imagery sampling format for 1280x720 progressive-scan High Definition (HD) Motion Imagery shall be defined by SMPTE ST 296.
MISP-2015.1-21	The Motion Imagery sampling format for 1920x1080 progressive-scan High Definition (HD) Motion Imagery shall be defined by SMPTE ST 274.

3.5.1.5.3 Ultra-High Definition

Requirement	
MISP-2015.1-22	The Motion Imagery sampling structures for 3840x2160 and 7680x4320 progressive-scan Ultra High Definition (UHD) Motion Imagery shall be defined by SMPTE ST 2036-1.

3.5.2 Containers

3.5.2.1 Serial Digital Interface (SDI)

A standardized family of commercial containers, known as Serial Digital Interface (SDI), provides transport for Class 0 Motion Imagery. These standards specify the data format and interface for communication between system components, and the data format for carrying a composition of various data types. Different family members support different data capacities (i.e., ED, HD, 4K, etc.). All provide additional data space for the carriage of non-imagery data. An Ancillary (ANC) packet (SMPTE ST 291 [24]) encapsulates any non-image data. The

Horizontal Ancillary Space (HANC) is the primary space for audio. The Vertical Ancillary Space (VANC) is available for basic data and is the space MISB allocates for KLV metadata.

MISB ST 0605 [25] provides guidance for inclusion of timestamps, metadata and audio in SDI.

3.5.2.2 GigE Vision

GigE Vision [26] is a communication interface based on Ethernet technology. It supports interfacing between a GigE Vision device and a network card using standard CAT-5e/6 cable, or any other physical medium supported by Ethernet. The GigE Vision specification defines device discovery; GigE Vision Control Protocol (GVCP), which is an application layer protocol allowing an application to configure a device (typically a camera) and to instantiate stream channels; and GigE Vision Streaming Protocol (GVSP), which is an application layer protocol allowing a GVSP receiver to receive image data, image information or other information from a GVSP transmitter.

MISB ST 1608 [27] is a profile of the GigE Vision specification and includes guidance for carriage of KLV metadata. ST 1608 specifies the use of MISB ST 1507 for providing timing information for imagery.

3.5.2.3 Material Exchange Format (MXF)

The Material Exchange Format (MXF), SMPTE ST 377-1 [28], finds use in applications such as the exchange of Motion Imagery Data amongst collection platforms, and in support of random access to databases based on metadata indexing. MXF provides advanced features for accessing and exchanging Motion Imagery Data over communication networks.

MISB ST 1606 [29] defines a file format for use in high performance and metric imaging applications within the DoD/IC/NSG. It mandates MXF with constraints for capturing, storing, exchange, play out, analysis, and archiving of Motion Imagery. Specifically designed to meet the high precision and accuracy requirements of Major Range and Test Facility Base (MRTFB) applications, this file format supports a wide range of pixel densities, frame rates, bit depths, and color formats, as well as very rich and detailed metadata.

Requirement	
MISP-2021.1-121	Class 0 Motion Imagery encapsulated in MXF shall comply with MISB ST 1606.

3.6 Class 1 Motion Imagery

Class 1 Motion Imagery represents compressed Motion Imagery for sampling formats typically found in commercial and consumer video applications.

3.6.1 Sampling Format

To maximize interoperability and data reuse Class 1 Motion Imagery spans a common set of sampling formats called Points of Interoperability (POI) derived from industry standard formats. For example, Table 3 for Class 0 Motion Imagery shows the 720p and 1080p formats foundational to high-definition television and 3840p format to ultra-high-definition television. Subsampling these formats by integer factors produces aspect-ratio-equivalent lower density formats.

3.6.2 Pixel Value Range

Class 1 Motion Imagery is typically composed of three (i.e., RGB) bands for visible light, and one band for infrared (i.e., monochrome) imagers. Each band has a maximum pixel value range of eight bits. Section 2.1.4.2 describes the more common 4:2:2 and 4:2:0 color models. Although there is perceptual color loss in these models, limiting color data is an initial step towards reducing data presented for compression. Both pixel value range and color model factor into the compression profile (discussed in the following sections) selected.

3.6.3 Compression

3.6.3.1 General

The MISP approves the following compression technologies for Class 1 Motion Imagery: H.265/HEVC (High Efficiency Video Encoding) [30] [31], H.264/AVC [32] [33] and H.262/MPEG-2 [34] [35]. H.265/HEVC is displacing H.264/AVC as a preferred compression technology because of its nearly 2:1 coding efficiency over H.264; this means a 50% reduction in data rate for a given level of image quality, or higher image quality at the same data rate. <u>The MISP recommends that systems produce H.265/HEVC</u>.

NOTE: Be advised that data compressed with H.265 may cause interoperability issues with DoD/IC/NSG architectures as many systems that decode Motion Imagery may not support H.265 currently. Any implementation that chooses to use H.265 should ensure its intended customers can receive/decode H.265 compressed Motion Imagery.

Compression *profiles* and *levels* serve as points of common interoperability between conformant implementations. Profiles are subsets of the bit-stream syntax that limit the algorithm options used when encoding video. Different profiles find use in different applications. Within the boundaries established by a profile there may still be wide variation in the computing resources required by encoders and decoders as frame size and frame rate changes. Levels place constraints on the memory and processing throughput required during encoding and decoding a bit-stream. These constraints may be simple limits on values. Alternatively, they may take the form of constraints on arithmetic combinations of values (e.g., picture width multiplied by picture height multiplied by number of pictures decoded per second).

A POI table for each compression type lists a set of sampling formats along with compression levels defined in the corresponding compression standard. The POI's are examples of common scanning formats typically found. An encoder is not bound to use these formats.

In the POI tables, for sampling formats where images per second may be less than the native rate, skipping images ensures smooth playback. For example, given a native rate of 60 images per second and a display rate of 30 images per second skip one image for every two; for a rate of 20 images per second skip two images for every three; for a rate of 15 images per second skip three images for every four, etc.

3.6.3.2 H.265/HEVC

H.265/HEVC is the next generation successor in the MPEG compression technology family with a compression efficiency roughly 2-to-1 over H.264/AVC at the same perceived quality level and supports increased pixel value range, spatial and temporal sampling formats.

Table 5 lists the POI for H.265/HEVC. Columns x Rows specifies the number of pixels horizontally and vertically per image. Images per Second indicates the image rate or range of image rates. H.265 Level corresponds to the level specified in the H.265/HEVC standard [30] [31] which supports the compression of Motion Imagery for the sampling format listed.

Table 5: Exemplar Points of Interoperability (H.265)

Sampling Format		H.265 Level
Columns x Rows	**Images per Second**	
3840 x 2160	60	5.1
	30 and less	5
1920 x 1080	60	4.1
	30 and less	4
1280 x 720	60	4
	30 and less	3.1
960 x 540	60	3.1
	30 and less	3
640 x 480	60	3.1
	30 and less	3
640 x 360	60	3
	30 and less	2.1
480 x 270	60	3
	30 and less	2.1
320 x 240	60	2.1
	30 and less	2
320 x 180	60	2

Version 1 of H.265/HEVC specifies three profiles: Main, Main 10 and Main Still Picture. Version 2 adds 21 range extension profiles, two scalable extension profiles, and one multi-view profile. Within the bounds of a profile HEVC allows a large variation in performance depending on selected bitstream elements. HEVC introduces two tiers within a profile to meet application needs. The Main tier applies to most applications and is the only option for levels below Level 4. The High tier supports very demanding high-bitrate applications; the High tier less likely applies to Class 1 Motion Imagery.

Requirement(s)	
MISP-2018.1-99	While compressing Class 1 Motion Imagery with H.265/HEVC, the compression shall comply with ISO/IEC 23008-2 | ITU-T Rec. H.265.
MISP-2018.2-113	While compressing Class 1 Motion Imagery with H.265/HEVC, the compression shall be profile Main 10 in the range of Level 1 to Level 5.1 inclusive.

3.6.3.3 H.264/AVC

Table 6 lists the POI for H.264/AVC. H.264 Level corresponds to the level specified in the H.264/AVC standard [32] [33] which supports the compression of Motion Imagery for the sampling format listed.

Table 6: Exemplar Points of Interoperability (H.264)

Sampling Format		H.264 Level
Columns x Rows	**Images per Second**	
1920 x 1080	30 and less	4
1280 x 720	60	3.2
	30 and less	3.1
960 x 540	60	3.2
	30 and less	3.1
640 x 480	60	3.1
	30	3
	15 and less	2.2
640 x 360	60	3.1
	30	3
	15 and less	2.2
480 x 270	60	3
	30 and less	2.1
320 x 240	60	2.1
	30	1.3
	15	1.2
	10 and less	1.1
320 x 180	60	2.1
	30	1.3
	15	1.2
	10 and less	1.1

H.264/AVC specifies thirteen profiles; common ones include Baseline, Constrained Baseline, Main and High. MISB RP 0802 [36] provides additional information on choices in profiles and levels. Consult MISB TRM 1404 [6] for further information on H.264/AVC compression including profiles, levels, and application-specific uses.

Requirement(s)	
MISP-2015.1-32	While compressing Class 1 Motion Imagery with H.264/AVC, the compression shall comply with ISO/IEC 14496-10 | ITU-T Rec. H.264.
MISP-2018.2-114	While compressing Class 1 Motion Imagery with H.264/AVC, the compression shall be profile Constrained Baseline, Main or High in the range of Level 1 to Level 4 inclusive.

3.6.3.4 Legacy Systems

Table 7, lists the POI for legacy systems. These systems typically produce compressed Motion Imagery using sampling formats which may be interlaced-scan, Standard Definition, and derivative formats compressed with either H.264/AVC or H.262/MPEG-2 compression.

Table 7: Exemplar Points of Interoperability for Legacy Systems

Sampling Format		H.264 Level (minimum)	MPEG-2 Level
Columns x Rows	Images per Second		
720 x 480	30	3	MP@ML
	15 and less	2.2	
640 x 480	30 and less	See Table 6	
352 x 240	30	1.3	
	15	1.2	
	5	1.1	
320 x 240	30 and less	See Table 6	
176 x 120	30	1.1	N/A
	15	1	

The 720x480 and 640x480 sampling formats are common commercial interlaced-scan formats found in legacy systems (although interlaced in not permitted in the MISP).

The H.264 Level and MPEG-2 columns indicate the H.264/AVC levels and H.262/MPEG-2 profiles and levels supporting a sampling format. MPEG-2 profiles and levels carry different designations than H.264/AVC. The H.262/MPEG-2 compression standard [34] identifies five profiles, with Main Profile (MP) and High Profile (HP) the most common. The coding description in MPEG-2 is a combination of profile and level as Profile@Level. For example, MPEG-2 MP@ML means MPEG-2 compression using Main Profile and Main Level.

Requirement(s)	
MISP-2015.1-34	While compressing Class 1 Motion Imagery with H.262/MPEG-2, the compression shall comply with ISO/IEC 13818-2 | ITU-T Rec H.262.
MISP-2018.2-115	While compressing Class 1 Motion Imagery with H.262/MPEG-2, the compression profile shall be Main at Main or High Level.

3.6.3.5 Infrared Systems

Infrared Motion Imagery may have a pixel value range which exceeds the limits of Class 1 Motion Imagery. MISB ST 0404 [23] provides guidance and recommendations in applying several types of mapping for reducing the pixel value range to meet the limits of Class 1 Motion Imagery.

Requirement	
MISP-2015.1-37	When infrared Motion Imagery with a pixel value range greater than 8 bits is converted into Class 1 Motion Imagery and compressed using H.262/MPEG-2 or H.264/AVC the compressed imagery shall comply with MISB ST 0404.

3.6.4 Decompression

3.6.4.1 General

Decoders which comply with H.265/HEVC, H.264/AVC and H.262/MPEG-2 compression must operate at a profile and level (as discussed in Sections 3.6.3.2, 3.6.3.3 and 3.6.3.4). A decoder cannot claim conformance to a profile and only partially support the bit-stream syntax of that profile. Likewise, a decoder cannot claim conformance to a level and only partially support the constraints on values of the syntax elements in the bitstream of that level. Conformance requirements for decoders encompass a broader class of bitstreams than those for encoders.

For example, an encoder conformant to a H.264/AVC profile and level need only produce <u>one</u> bitstream conformant to that profile and level; a conformant decoder, however, must be able to decode <u>*all*</u> bitstreams conformant to their profile and level. As discussed in Sections 3.6.3.2 and 3.6.3.3, the POI shown in Table 5 apply to H.265/HEVC compressed bitstreams, while those in Table 6 apply to H.264/AVC compressed bitstreams. For an encoder which produces a H.264/AVC Level 3.2 bitstream, a decoder must be able to decode that same H.264/AVC Level 3.2 bitstream. In addition, that decoder must be able to decode all H.264/AVC bitstreams produced by encoders below Level 3.2 (i.e., Level 3.1, 3.0, etc.).

A MISP conformant decoder needs to decode H.265/HEVC, H.264/AVC or H.262/MPEG-2 at the appropriate profile and level designated for that specific sampling format. In addition, it needs to support the decoding of all levels below that level. Systems that use non-conformant decoders may be able to support one or more of the POI, thus enabling some measure of interoperability with MISP conformant systems.

ISO/IEC 13818-2 [34] | ITU-T Rec H.262 [35] governs decoder conformance requirements for H.262/MPEG-2. ISO/IEC 14996-10 [32] | ITU-T Rec. H.264 [33] governs decoder conformance requirements for H.264/AVC. ISO/IEC 23008-2 [30] | ITU-T Rec. H.265 [31] governs decoder conformance requirements for H.265/HEVC.

Requirement(s)	
MISP-2015.1-38	A Class 1 Motion Imagery decoder shall support the decoding of Class 1 Motion Imagery compressed using H.262/MPEG-2 Main Profile at Main and High Level.

MISP-2015.1-39	A Class 1 Motion Imagery decoder for H.262/MPEG-2 shall fully meet the conformance requirements of ISO/IEC 13818-2 | ITU-T Rec H.262 per profile and level.
MISP-2017.1-94	A Class 1 Motion Imagery decoder shall support the decoding of Class 1 Motion Imagery compressed using H.264/AVC Constrained Baseline Profile, Main Profile and High Profile at Level 4.
MISP-2015.1-41	A Class 1 Motion Imagery decoder for H.264/AVC shall fully meet the conformance requirements of ISO/IEC 14496-10 | ITU-T Rec. H.264 per profile and level.
MISP-2018.1-102	A Class 1 Motion Imagery decoder shall support the decoding of Class 1 Motion Imagery compressed using H.265/HEVC profile Main 10 at Level 5.1.
MISP-2018.1-103	A Class 1 Motion Imagery decoder for H.265/HEVC shall fully meet the conformance requirements of ISO/IEC 23008-2 | ITU-T Rec. H.265 per profile and level.

3.6.5 Annotations

Not all decoders support annotations. The MISB provides guidance for adding annotations in a non-destructive manner to Motion Imagery (i.e., new information does not overwrite original image pixels). MISB ST 0602 [37] provides direction on the creation of "Annotation" KLV metadata to allow for the creation, dissemination, and display of visual cues to enhance the exploitation of Motion Imagery Data. The following requirement applies to a decoder that supports annotations.

Requirement	
MISP-2015.1-42	Where a Class 1 Motion Imagery decoder supports graphic overlay, the decoder shall comply with MISB ST 0602.

3.6.6 Timestamps

A timestamp based on Absolute Time as defined in MISB ST 0603 [12] is mandatory in Class 1 Motion Imagery within each image independent of the metadata. H.265/HEVC, H.264/AVC and H.262/MPEG-2 define specific locations in their respective compressed bitstreams for this information. MISB ST 0604 [38] provides guidance on which fields in the bitstreams to insert the timestamp, as well as a defined format for the information.

Requirement	
MISP-2018.1-104	Class 1 Motion Imagery shall contain a timestamp based on Absolute Time in accordance with MISB ST 0604.

3.6.7 Metadata

The purpose for Class 1 Motion Imagery is to meet constrained bitrate criteria for delivery. Adding extra data such as metadata and/or audio can reduce the bit-space available for the imagery, hence potentially reducing the subjective quality of the Motion Imagery. For this reason, Key Length Value (KLV), which is an extremely bit-efficient method for conveying information, is the encoding method for the metadata.

Requirement	
MISP-2015.1-45	Class 1 Motion Imagery metadata shall be represented using KLV (Key Length Value).

The Motion Imagery Handbook defines common metadata types and describes how to organize the sensor/platform data into a hierarchy of KLV Packs and Local Sets that reduces the bandwidth needed to transmit the data.

3.6.8 Audio

The MISB has approved several compressed formats (see MISB ST 1001 [39]) intended to facilitate the interoperable use of audio.

Requirement	
MISP-2015.1-46	Class 1 Motion Imagery compressed audio shall comply with MISB ST 1001.

3.6.9 Containers

The following subsections specify the requirements for the container type and structure for Class 1 Motion Imagery.

3.6.9.1 MPEG-2 Transport Stream

The MPEG-2 Transport Stream (TS) container [40] [41] [42] provides for the carriage of Motion Imagery, metadata and audio as a unified package, as shown in Figure 9. Internal timing signatures for the Motion Imagery, metadata (when in the synchronous channel), and audio enable the data to retain relationships for presentation at the display. MISB TRM 0909 [43] provides additional information on constructing a MPEG-2 Transport Stream that complies with the MISP.

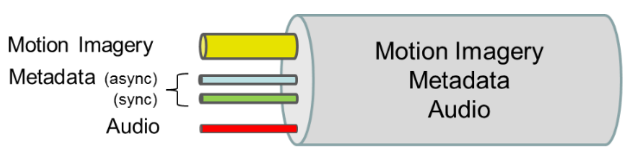

Figure 9: MPEG-2 Transport Stream Container

MISB ST 1402 [44] profiles the MPEG-2 TS reflected in the following requirements.

Requirement(s)	
MISP-2015.1-47	Class 1 Motion Imagery encapsulated in a MPEG-2 Transport Stream container shall comply with ISO/IEC 13818-1 | ITU-T Rec H.222.0.
MISP-2015.1-48	Class 1 Motion Imagery encapsulated in a MPEG-2 Transport Stream container shall comply with MISB ST 1402.
MISP-2015.1-49	Security metadata encapsulated in a MPEG-2 Transport Stream container shall be inserted into only one of the two carriage mechanisms available: The Synchronous Stream Multiplex Method or the Asynchronous Stream Multiplex Method in accordance with MISB ST 1402.

MISP-2015.1-50	Class 1 Motion Imagery encapsulated in a MPEG-2 Transport Stream container shall meet the conformance requirements of ISO/IEC 13818-4.

3.6.9.2 Common Media Application Format (CMAF)

The Common Media Application Format (CMAF) ISO/IEC 23000-19 [45] container provides for encoding and packaging segmented media – Motion Imagery, audio, and metadata – for delivery via server/client-based protocols such as MPEG-DASH [46].

CMAF usage as defined by MISB ST 1910 [47] supports H.264/AVC and H.265/HEVC compressed Motion Imagery, compressed audio, and KLV metadata.

Requirement(s)	
MISP-2021.1-122	Class 1 Motion Imagery encapsulated in a CMAF container shall comply with ISO/IEC 23000-19.
MISP-2021.1-123	Class 1 Motion Imagery encapsulated in a CMAF container shall comply with MISB ST 1910.
MISP-2021.1-124	Security metadata encapsulated in a CMAF container shall be present in either the synchronous or asynchronous metadata message boxes, but not both.

3.7 Class 2 Motion Imagery

Class 2 Motion Imagery represents compressed Motion Imagery, where one or more of the image characteristics (see Section 2.1.4.1) exceeds the limits of Class 1 Motion Imagery. An example is Large Volume Motion Imagery (LVMI), where a variety of image geometries ranging from a single image to a matrix of individual images composited into one complete single image.

Figure 10 illustrates a 2x2 array, four-camera, multi-image imager that produces one complete image, where different imager geometries (e.g., linear, larger array, etc.) are possible.

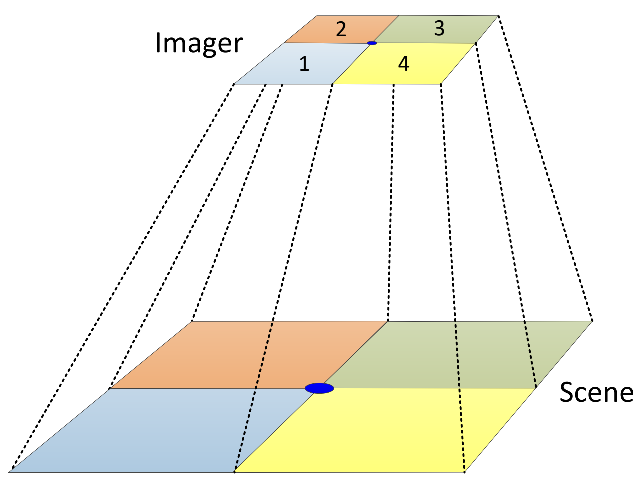

Figure 10: Illustration of 2x2 Image Geometry

Large datasets are a byproduct of LVMI. Because of the quantity of data collected per image and/or collected over extended time intervals, the large data-handling requirements, such as storage and dissemination, incur significant costs. LVMI systems may operate at image rates outside the bounds of Class 1 Motion Imagery, which further contributes to dissemination and exploitation issues.

Another Class 2 Motion Imagery application is Scientific/Engineering Motion Imagery (SEMI), where for example the Motion Imagery sample value range may be more than 10 bits per band, there may be more than three color/spectral bands, or the images per second may be in the hundreds or thousands. There may also be a need for "perceptually lossless" compression, where the Motion Imagery compression does not inject artifacts detectable by human observation.

3.7.1 Sampling Format

Class 2 Motion Imagery may have characteristics, such as sample value range, pixel value range, image size, image aspect ratio, bands per image and number of images per second representative of Class 0 Motion Imagery but are outside those of Class 1 Motion Imagery.

3.7.2 Pixel Value Range

The pixel value range for Class 2 Motion Imagery varies from 12 bits up to 32 bits.

3.7.3 Image Size and Aspect Ratios

The image size of Class 2 Motion Imagery and its aspect ratio can vary greatly. For example, a linear array of three 1920x1080 HD cameras could create a single image with 5760 columns and 1080 rows resulting in a 16:3 aspect ratio.

3.7.4 Bands

Class 2 Motion Imagery may have any number of bands depending on their intended spectral coverage. For example, a HSI image could require hundreds of bands.

3.7.5 Number of Images per second

Class 2 Motion Imagery may have any number of images per second depending on the application.

3.7.6 Compression

The MISB recommends JPEG2000 [48], H.265/HEVC and H.264/AVC compression for Class 2 Motion Imagery. JPEG2000 finds use in the compression of LVMI datasets. JPEG2000 also finds use in other Class 2 Motion Imagery applications, where other technologies like H.265/HEVC and H.264/AVC are less appropriate. JPEG2000 compression has different profiles. To maximize interoperability across Class 2 Motion Imagery Systems that use JPEG2000, the following requirement ensures the use of a specific profile.

Requirement	
MISP-2015.1-51	Class 2 Motion Imagery compressed using JPEG2000 shall comply with ISO/IEC BIFF Profile BPJ2K.

In applications where the Motion Imagery has a pixel value range that exceeds that of Class 1 Motion Imagery, H.264/AVC can accommodate a pixel value range up to 14 bits using more advanced profiles of H.264/AVC. H.265/HEVC can accommodate a pixel value range up to 16 bits using more advanced profiles of H.265/HEVC.

Within the bounds of a profile HEVC allows a large variation in performance depending on selected bitstream elements. HEVC introduces two tiers within a profile to meet application needs. The Main tier applies to most applications and is the only option for levels below Level 4. The H.265 High tier supports very demanding applications and may apply in some Class 2 Motion Imagery use cases. Users should be aware that these are uncommon profiles potentially requiring special decoders.

Requirement(s)	
MISP-2018.3-117	While compressing Class 2 Motion Imagery with H.264/AVC the compression shall be High 4:2:2 Profile (Hi422P) or High 4:4:4 Predictive Profile (Hi444PP) in the range of Level 1 to Level 4.2 inclusive.
MISP-2015.1-53	Infrared Class 2 Motion Imagery compressed with H.264/AVC shall comply with MISB ST 0404.
MISP-2018.3-118	While compressing Class 2 Motion Imagery with H.265/HEVC the compression shall be Main Profile 4:2:2 12 or Main Profile 4:4:4 12 in the range of Level 1 to Level 6.1 inclusive.

3.7.7 Decompression

A MISP conformant decoder needs to decode at the appropriate profile and level designated for that specific sampling format. In addition, it needs to support the decoding of all levels below that level. Systems that use non-conformant decoders may be able to support one or more of the POI, thus enabling some measure of interoperability with MISP conformant systems.

ISO/IEC 15444-4 [49] governs decoder conformance requirements for JPEG2000. ISO/IEC 14996-10 [32] | ITU-T Rec. H.264 [33] governs decoder conformance requirements for H.264/AVC. ISO/IEC 23008-2 [30] | ITU-T Rec. H.265 [31] governs decoder conformance requirements for H.265/HEVC.

Requirement(s)	
MISP-2015.1-54	A Class 2 Motion Imagery decoder for JPEG2000 shall meet the conformance requirements of ISO/IEC 15444-4.
MISP-2015.1-55	A Class 2 Motion Imagery decoder for H.264/AVC shall support the decoding of High 4:2:2 Profile (Hi422P) and High 4:4:4 Predictive Profile (Hi444PP) at Level 4.2.
MISP-2015.1-56	A Class 2 Motion Imagery decoder for H.264/AVC shall fully meet the conformance requirements of H.264/AVC per profile and level.
MISP-2018.3-119	A Class 2 Motion Imagery decoder for H.265/HEVC shall support the decoding of the profiles Main 4:2:2 12 and Main 4:4:4 12 at Level 6.1.
MISP-2018.1-107	A Class 2 Motion Imagery decoder for H.265/HEVC shall fully meet the conformance requirements of H.265/HEVC per profile and level.

3.7.8 Annotations

The requirements for annotations for Class 2 Motion Imagery encapsulated within a MPEG-2 Transport Stream are like those of Class 1 Motion Imagery (see Section 3.6.5).

Requirement	
MISP-2016.1-85	Where a Class 2 Motion Imagery decoder supports graphic overlay, the decoder shall comply with MISB ST 0602.

MIL-STD-2500 [17]/STANAG 4545 [18] governs the use of annotations for Class 2 Motion Imagery encapsulated within a Motion Imagery Extensions for NITF (MIE4NITF [50]) container.

3.7.9 Timestamps

The requirements for insertion of a timestamp into Class 2 Motion Imagery compressed using H.265/HEVC and H.264/AVC are like those of Class 1 Motion Imagery (see Section 3.6.6).

Requirement	
MISP-2018.1-108	Class 2 Motion Imagery shall contain a timestamp based on Absolute Time in accordance with MISB ST 0604.

The MIE4NITF [50] specification governs the insertion of a timestamp into Class 2 Motion Imagery encapsulated within a MIE4NITF container.

3.7.10 Metadata

When encapsulating Class 2 Motion Imagery and metadata into a MPEG-2 Transport Stream, the Class 1 Motion Imagery metadata standards and requirements apply.

When encapsulating Class 2 Motion Imagery and metadata within a MIE4NITF [50] container, the requirements for metadata are defined by MIL-STD-2500 [17]/STANAG 4545 [18] and/or program implementation-specific profiles.

3.7.11 Audio

The requirements for audio in Class 2 Motion Imagery encapsulated within a MPEG-2 Transport Stream are like those for Class 1 Motion Imagery (see Section 3.6.8).

Requirement	
MISP-2016.1-88	Class 2 Motion Imagery compressed audio shall comply with MISB ST 1001.

3.7.12 Containers

3.7.12.1 MPEG-2 Transport Stream

The requirements listed below are specific to MPEG-2 TS as a container for H.264/AVC and H.265/HEVC compressed Class 2 Motion Imagery; however, any non-specific class requirements listed for Class 1 Motion Imagery (see Section 3.6.9.1) are likewise mandatory.

Requirement(s)	
MISP-2016.1-89	Class 2 Motion Imagery encapsulated in a MPEG-2 Transport Stream container shall comply with ISO/IEC 13818-1 | ITU-T Rec H.222.0.
MISP-2016.1-90	Class 2 Motion Imagery encapsulated in a MPEG-2 Transport Stream container shall comply with MISB ST 1402.
MISP-2016.1-91	Class 2 Motion Imagery encapsulated in a MPEG-2 Transport Stream container shall meet the conformance requirements of ISO/IEC 13818-4.

3.7.12.2 MIE4NITF

The MIE4NITF [50] container is an extension to MIL-STD-2500 [17]/STANAG 4545 [18] for Class 2 Motion Imagery applications that require spatial blocking, greater than three spectral bands, greater than 10-bit pixel value range, etc., such as LVMI and SEMI.

Requirement	
MISP-2015.1-57	Class 2 Motion Imagery placed into a NITF (National Imagery Transmission Format)/NSIF (NATO Secondary Imagery Format) shall comply with MIE4NITF.

3.7.12.3 Material Exchange Format (MXF)

MISB ST 1606 defines a file format for use in high performance and metric imaging applications within the DoD/IC/NSG. It mandates MXF with constraints for capturing, storing, exchange, play out, analysis, and archiving of Motion Imagery. Specifically designed to meet the high precision and accuracy requirements of MRTFB applications, this file format supports a wide range of pixel densities, frame rates, bit depths, and color formats, as well as very rich and detailed metadata.

Requirement	
MISP-2021.1-125	Class 2 Motion Imagery encapsulated in MXF shall comply with MISB ST 1606.

3.7.12.4 Common Media Application Format (CMAF)

The Common Media Application Format (CMAF) ISO/IEC 23000-19 container provides for encoding and packaging segmented media – Motion Imagery, audio, and metadata – for delivery via server/client-based protocols such as MPEG-DASH.

CMAF usage as defined by MISB ST 1910 supports H.264/AVC and H.265/HEVC compressed Motion Imagery, compressed audio, and KLV metadata. The requirements listed below are specific to CMAF as a container for H.264/AVC and H.265/HEVC compressed Class 2 Motion Imagery; however, any non-specific class requirements listed for Class 1 Motion Imagery (see Section 3.6.9.2) are likewise mandatory.

Requirement(s)	
MISP-2021.1-126	Class 2 Motion Imagery encapsulated in a CMAF container shall comply with ISO/IEC 23000-19.
MISP-2021.1-127	Class 2 Motion Imagery encapsulated in a CMAF container shall comply with MISB ST 1910.

3.8 Class 3 Motion Imagery

Class 3 Motion Imagery applies to Motion Imagery Systems which produce content outside the MISP ecosystem of standards and requirements. Class 3 Motion Imagery includes those sources external to the NSG, typified by content from cell phones, mobile devices, surveillance cameras, and other industrial cameras. These sources of content come in various compressed video/audio formats and containers that may not comply with the MISP. As these sources represent potential intelligence information, reformatting the content to MISP standards enables use in exploitation tools.

3.8.1 Motion Imagery Data Normalization

Content normalization is the process of transforming Class 3 Motion Imagery to comply with MISP standards and requirements. These issues include:

- Non-supported Motion Imagery compression
- Lack of or insufficient metadata (i.e., none or time and location only)
- Non-supported metadata encodings
- Non-supported audio compression
- Non-supported container

The following process (shown in Figure 11) normalizes content to MISP Standards:

- Store the original source content (if warranted)
- Remove all content from its container
- Decompress and compress data to comply with the MISP formats (if necessary)
- Insert a timestamp based on Absolute Time into the Motion Imagery
- Add metadata to include source ID, timestamp, location information
- Encapsulate into an MISP-approved container (i.e., MPEG-2 TS, etc.)

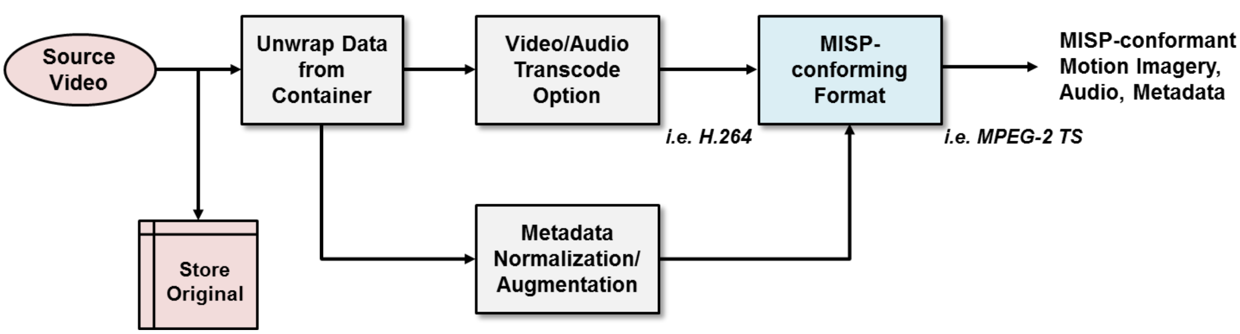

Figure 11: Motion Imagery Data Normalization

The following requirements normalize non-conforming MISP Motion Imagery, metadata, and audio for exploitation purposes.

Requirement(s)	
MISP-2018.1-109	Class 3 Motion Imagery shall be compressed using H.265/HEVC, H.264/AVC or H.262/MPEG-2.

MISP-2018.1-110	When timestamp information is available in Class 3 Motion Imagery, it shall be converted to a timestamp based on Absolute Time in accordance with MISB ST 0603.
MISP-2018.1-111	When a timestamp based on Absolute Time is generated from Class 3 Motion Imagery, it shall be inserted into a H.265/HEVC, H.264/AVC or H.262/MPEG-2 compressed elementary stream in accordance with MISB ST 0604.
MISP-2018.1-112	When timestamp information is available in Class 3 Motion Imagery and after conversion to a timestamp based on Absolute Time in accordance with MISB ST 0603, it shall be inserted as metadata in accordance with MISB ST 0601.
MISP-2015.1-62	When location information is available in Class 3 Motion Imagery, it shall be inserted as metadata in accordance with MISB ST 0601.
MISP-2015.1-63	When converting Class 3 Motion Imagery to meet MISP requirements, a Minor Core Identifier shall be generated and inserted in accordance with MISB ST 1204.
MISP-2015.1-64	Class 3 Motion Imagery audio shall be compressed in accordance with MISB ST 1001.
MISP-2015.1-65	When converting Class 3 Motion Imagery to meet MISP requirements, the mandatory security metadata items from MISB ST 0102 shall be included.
MISP-2015.1-66	When converting Class 3 Motion Imagery to meet MISP requirements, it shall be encapsulated in a MISP-approved container.

4 CONTENT

4.1 Introduction

This section provides requirements for the content that maps to the structure presented in Section 3. The content is comprised of three types: Motion Imagery, audio and metadata. The Motion Imagery is the visual information exploited by analysts. The types of content represented in Motion Imagery are vast and dependent on the scene type, sensor position, modality, and other factors. While it is difficult to levy requirements on the scene data, the MISP does provide standards about the quality of the Motion Imagery.

The audio is optional content included with the Motion Imagery. The source of the audio can be from the sensor itself or added to the Motion Imagery after initial capture downstream in exploitation.

Metadata provides the information for contextual interpretability (see Section 2.1.2). Motion Imagery by itself has limited intelligence value without context. Providing as much contextual information about the Motion Imagery both when collected, and later in exploitation, greatly increases its intelligence and tactical value. While structures built on standards and requirements ensure a common framework for data sharing and analysis, the content placed into those structures is most important for exploitation. The quality of the metadata – both its availability and accuracy - is critical. Ongoing improvement in data collection is continuous and must be a primary focus within the community of practice.

4.2 Motion Imagery - Content

Achieving consistency in Motion Imagery content requires measuring and monitoring Motion Imagery quality. The atmosphere, optics, compression, delivery methods and other factors all affect Motion Imagery quality on its way to exploitation. The goal is to produce the highest quality Motion Imagery content, then maintain this level of quality. This begins with the Motion Imagery itself and its management throughout the process of capture, delivery, and display. Metadata in the Quality domain (see Section 4.4.2.8) aides in monitoring Motion Imagery quality.

Displaying text and graphics over Motion Imagery provides a visual aid for some users, however for other users and automated systems removing the graphics helps in exploiting the Motion Imagery; the following requirement addresses this issue.

Requirement	
MISP-2015.1-67	Where graphic and text information are overlaid onto Motion Imagery, the information shall be nondestructive to the Motion Imagery content (i.e., “burned-in metadata” is not allowed).

MISB ST 1909 [51] identifies metadata content with locations that adjust according to display resolution to create a nondestructive, common metadata overlay for Motion Imagery visualization.

4.3 Audio Content

At present, the MISP does not provide standards for audio content other than the quality loss discussion in MISB ST 1001 [39].

4.4 Metadata Content

Many metadata items provide contextual interpretability and other important information for Motion Imagery. Domains facilitate organizing metadata items. Applications draw from several of these domains to meet their requirements. A collection is a group of domains selected to support an application or use case.

Applications require different types of encoding for metadata and the type can be different within a class of Motion Imagery. Class 1 Motion Imagery mandates encoding all metadata as KLV – this is to ensure the most efficient use of bits within constrained channels. The encoding of metadata for Class 2 Motion Imagery depends on the container. For example, Class 2 Motion Imagery encapsulated in MPEG-2 Transport Stream and MXF likewise encode metadata as KLV, whereas in MIE4NITF metadata encodes as Tagged Record Extensions (TRE).

In all uses of metadata sets governed by the MISB all mandatory requirements indicated within a metadata standard are to be satisfied.

Requirement	
MISP-2016.1-92	A MISB metadata set shall conform to all requirements as specified for that metadata set.

4.4.1 Domains and Collections Overview

Domains are metadata items related around a central topic. For example, the Temporal domain is a group of information about time; it includes Absolute Time, relative time, temporal accuracies, time sources, lock states, etc. Aside from mandated items, an application may choose items from a domain where needed, for example, Absolute Time, whereas other applications may use them all. Domains generally have specified requirements. Certain metadata items are mandatory within a given domain. Depending on the Motion Imagery class a domain may have different requirements.

Domains can be related to other domains; the relationship can be hierarchical or a singular dependency (i.e., a graph). Example domains include: Administrative, Combat, Exploitation, Image Space, Infrared (IR), Platforms (Aeronautical, Ground, Marine, etc.), Quality, Synthetic Aperture Radar (SAR), Temporal, Visual Light, and others. Although domains are topic driven a goal is to keep the size of the domain (i.e., the number of metadata items in a domain) as small as possible. Thus, breaking a domain into smaller domains, if possible, makes sense when a domain becomes too large.

Collections are associations of domains, along with some usage rules, for a specific application. For example, MISB ST 0601 - UAS Datalink Local Set is a collection of metadata from several domains as shown in Figure 12.

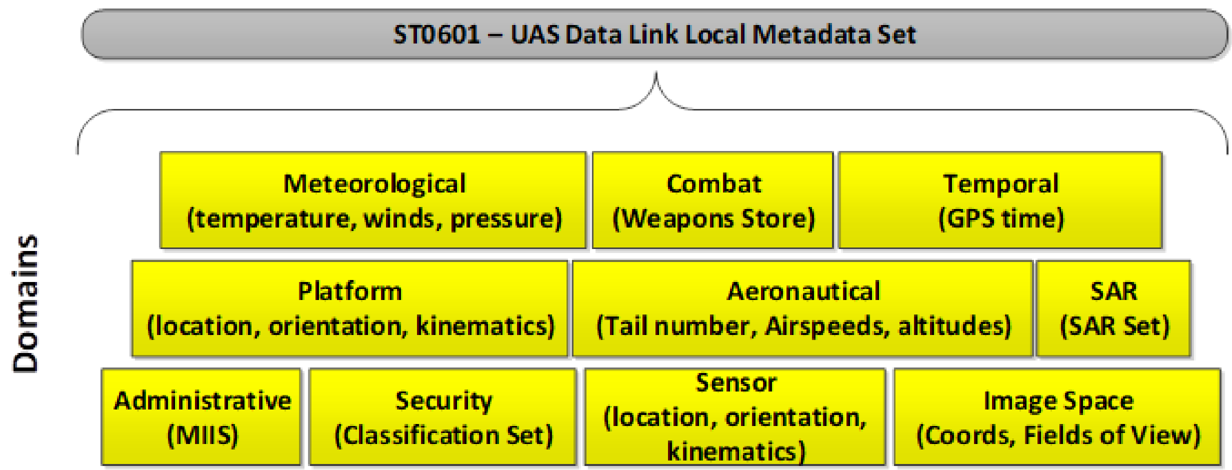

Figure 12: Example of a Collection (MISB ST 0601)

Collections leverage a common set of domains, and this commonality provides consistency among different collections, and therefore, interoperability amongst systems.

4.4.2 Metadata Domains

The concept of organizing metadata into domains and collections, first introduced in MISP-2015, improved reusability and document clarity. The following sections list some domains and their corresponding metadata.

4.4.2.1 Administrative Domain

The Administrative domain provides general metadata about the use and the management of Motion Imagery Data. An essential component of the Administrative domain is the ability to identify Motion Imagery. Many different sensors produce Motion Imagery distributed across many different networks and received by many different users and systems. Providing a consistent name or identity for each source enables the coordination of analysis and management with less confusion given many Motion Imagery sources. The Motion Imagery Identification System (MIIS) (MISB ST 1204 [52]) defines a mandatory and consistent unique identifier for all sensors and platforms. MISB ST 1301 [53] supports the MIIS with defined supplemental identifiers.

Requirement(s)	
MISP-2015.1-68	Motion Imagery shall contain a Core Identifier in accordance with MISB ST 1204.
MISP-2015.1-69	Where supplemental identifiers are used with MISB ST 1204, the supplemental identifiers shall be defined by MISB ST 1301.

4.4.2.2 Aeronautical Domain

The Aeronautical domain provides specific metadata about aerial platforms, such as aerostats, rockets, manned and unmanned aircraft. The Aeronautical domain is an extension to the Platform domain, so it only contains information that relates to aeronautical systems. Example information in the Aeronautical domain is tail number, wind speed, crabbing angle, etc.

MISB ST 0601 [54] contains items in the Aeronautical domain.

4.4.2.3 Combat Domain

The Combat domain provides metadata about weapons stores.

MISB ST 0601 [54] contains items in the Combat domain.

4.4.2.4 Exploitation Domain

The Exploitation domain is metadata created to assist Motion Imagery exploitation. The Exploitation domain is an extension to the Image Space domain. Exploitation tools and Motion Imagery pre-processing systems can generate Exploitation domain metadata; examples of tools are annotations, object detection, and tracking.

The Exploitation domain includes the following standards and mandated when providing the capabilities outlined in each of the following documents.

- MISB ST 0602 [37] provides direction on the creation of "annotation" KLV metadata to allow for the creation, dissemination, and display of visual cues to enhance the exploitation of Motion Imagery.
- MISB ST 0903 [55] defines KLV metadata to deliver Video Moving Target Indicator (VMTI) and Motion Imagery derived track metadata; the relationship between the VMTI metadata and other relevant standards; and gives implementation guidance for the VMTI metadata.
- MISB ST 0808 [56] documents KLV metadata supporting the creation, dissemination, and display of supplemental text to enhance the exploitation of Motion Imagery.
- MISB ST 0806 [57] defines KLV metadata to support a Remote Video Terminal (RVT).

4.4.2.5 Image Space Domain

The Image Space domain states or computes image space information. Image space information is either the relative or the absolute position (e.g., latitude/longitude/elevation) of each pixel in every image. Specifying this information for every pixel can be very bandwidth intensive, so estimates and models are employed to compute the image space information. Image space models can use platform dynamics (platform position and orientation) from the Platform domain and sensor dynamics from the Sensor domain to compute the pixel locations along with associated error estimates.

Image space information appears in several MISB documents. MISB ST 0601 [54] contains items that provide a threshold level of image space information. MISB ST 0801 [58] contains items that provide data for a photogrammetric level of image space information. MISB ST 1107 [59] specifies a threshold and objective profile of ST 0801.

Requirement	
MISP-2015.1-70	When implementing MISB ST 0801 the threshold and objective profiles for photogrammetric metadata items, the items shall be defined by MISB ST 1107.

The following requirements reference MISB ST 1202 [60] and MISB ST 1010 [61] which support image space processing.

Requirement(s)	
MISP-2015.1-71	Transforming two-dimensional Motion Imagery from one coordinate system into a second two-dimensional coordinate system shall comply with MISB ST 1202.
MISP-2015.1-72	Where Standard Deviation and Correlation Coefficient metadata is available, such metadata shall be provided in accordance with MISB ST 1010.

MISB ST 1002 [62] defines KLV metadata for the dissemination of an array of range data collected from the Standoff Precision Identification in Three-Dimensions (SPI-3D) Laser Detection And Ranging (LADAR) sensor.

4.4.2.6 Meteorological Domain

The Meteorological domain conveys atmospheric and weather metadata. MISB ST 0809 [63] defines the metadata and KLV structures for the Meteorological domain.

4.4.2.7 Platform Domain

The Platform domain provides the position, kinematics, and orientation metadata of the platform. MISB ST 0601 [54] contains items in the Platform domain.

4.4.2.8 Quality Domain

The Quality domain describes Motion Imagery quality at different points along the Functional Model beginning at the imager all the way through exploitation. Quality metadata serves two purposes: 1) it provides a measure of the health of the Motion Imagery at various points in delivery; and 2) it provides the exploiter with a measure of the "goodness" of the Motion Imagery.

- MISB ST 1108 [64] defines metadata for documenting image metrics and compression characteristics within a Motion Imagery stream or file. Such metadata facilitates documenting the quality of Motion Imagery along the imaging workflow with objective metrics or subjective interpretability criteria providing valuable information to the end user.

The Video National Imagery Interpretability Rating Scale (VNIIRS) is a user-based visual interpretability scale for Motion Imagery.

- MISB ST 0901 [65] includes the background of the VNIIRS design, instructions for its use, and rating criteria for Electro-Optical (EO) reflective sensors.

4.4.2.9 Security Domain

The Security domain provides metadata about the allowed access to Motion Imagery Data.

MISB ST 0102 [66] provides direction on the use of security metadata in Motion Imagery applications. Marking Motion Imagery Data correctly and consistently with security classification and other security administration information is mandatory. The approved practices in this standard are mandatory for all Motion Imagery Data.

Requirement	
MISP-2015.1-73	Motion Imagery shall include security metadata in accordance with MISB ST 0102.

4.4.2.10 Sensor Domain

The Sensor domain provides the position, kinematics, and orientation metadata of the sensor along with Sensor characteristics. Sensor domain information is described in several MISB documents; MISB ST 0601 [54] provides basic metadata and MISB ST 0801 [58] provides photogrammetric metadata.

4.4.2.11 Synthetic Aperture Radar (SAR) Domain

The Synthetic Aperture Radar (SAR) domain provides metadata about SAR Motion Imagery. MISB ST 1206 [67] defines the content and KLV metadata necessary to exploit both sequential SAR imagery and sequential SAR coherent change products as Motion Imagery. MISB ST 1403 [68] defines the threshold SAR metadata set.

Requirement	
MISP-2015.1-74	When implementing MISB ST 1206, the threshold profiles for Synthetic Aperture Radar Motion Imagery metadata items shall be defined by MISB ST 1403.

4.4.2.12 Temporal Domain

The Temporal domain includes metadata about time and its accuracy. This information relies on an absolute, reliable, common time reference, which is essential in timestamping Motion Imagery and metadata collected in operations. For this purpose, MISB ST 0603 [12] identifies two timestamps representing Absolute Time – the Precision Time Stamp and the Nano Precision Time Stamp.

MISB ST 1603 [69] defines metadata to indicate the lock and synchronization status of clocks within a timing system. Time transfer is the process of continuously synchronizing two or more clocks producing a hybrid time source. In time transfer, one clock is a reference clock, and the other clock(s) are receptor clocks. With Motion Imagery Systems, time transfer shares the reference time of a well-known accurate clock (e.g., GPS) with a local receptor clock (e.g., GPS receiver). Time transfer is a continuous process where each receptor clock is constantly comparing and adjusting its time to match the reference clock.

MISB ST 1507 specifies a metadata model for exposure configurations to represent timing information in global shutter and rolling shutter imagers.

4.4.2.13 Orbital Location Domain

The Orbital Location domain provides position information of an orbiting Motion Imagery collector. MISB ST 1504 [70] defines the content and KLV metadata necessary to communicate effectively, efficiently, and with the greatest resulting accuracy the information (the state vector) necessary to determine the position of an orbiting Motion Imagery collector.

4.4.3 Collections

Motion Imagery drives many different applications, and for each of these applications a collection defines the domain(s) needed to support an application.

4.4.3.1 Space Collection

The MISP currently does not specify a collection for Space applications; therefore, MISB recommends use of the Airborne collection to the extent possible.

4.4.3.2 Airborne Collection

The Airborne collection contains information and rules from several domains: Administrative, Aeronautical, Combat, Image Space, Platform, SAR, Sensor, Security and Temporal. Each domain provides requirements and rules of use.

It is mandatory for airborne platforms that produce Motion Imagery to incorporate the Motion Imagery Sensor Minimum Metadata Set (MISMMS), which enables the basic capabilities of situational awareness, discovery & retrieval, and cross-domain dissemination. MISB ST 0902 [71], which is a prerequisite for MISP conformance, defines the MISMMS as a profile of MISB ST 0601.

Requirement	
MISP-2015.1-75	Motion Imagery shall contain the Motion Imagery Sensor Minimum Metadata Set in accordance with MISB ST 0902.

4.4.3.3 Surface Collection

The first instantiation of the in-development Motion Imagery Metadata (MIMD) Model discussed in Section 4.4.5 addresses the Surface Collection. MISB ST 1801 [72] provides the baseline recommendation/profile for the Surface Collection.

4.4.3.3.1 Ground-Surface Collection

For the Ground-Surface collection the MISB recommends use of MISB ST 1801, and in collaboration with the MISB, add model elements to ST 1801 where appropriate.

4.4.3.3.2 Water-Surface Collection

For the Water-Surface collection the MISB recommends use of MISB ST 1801, and in collaboration with the MISB, add model elements to ST 1801 where appropriate.

4.4.3.4 Underwater Collection

The first instantiation of the in-development MIMD Model discussed in Section 4.4.5 addresses the Surface collection. For the Underwater collection, the MISB recommends use of MISB ST 1801, and in collaboration with the MISB, add model elements to ST 1801 where appropriate.

4.4.4 Constructs to Amend/Segment KLV Metadata

MISB ST 1607 [73] defines two KLV metadata Local Sets which enable parent/child relations, where reuse of metadata items within an instantiating parent metadata set extend the capabilities of that metadata set for new applications. These Local Sets also provide a means to include additional application-specific metadata sets, wherein data items within such an application-specific metadata set are independent of an instantiating metadata set. In effect, the Segment LS and Amend LS inherit metadata items from their instantiating metadata set based on application need. When added to an instantiating MISB metadata set, these KLV constructs provide new functionality, such as metadata editing, preservation and sharing, thereby extending the instantiating metadata set to meet new application needs.

- MISB ST 1601 [74] leverages the Amend LS in ST 1607 and defines metadata to support the identification of a geo-registration algorithm and various outputs from a geo-registration process.
- MISB ST 1602 [75] leverages the Segment LS in ST 1607 to assign image-attribute metadata to several images composited into one image frame.

Requirement	
MISP-2017.1-95	When metadata items within an instantiating metadata set are changed, the changed metadata shall be signaled using the Amend Local Set and Segment Local Set as defined in MISB ST 1607.

4.4.5 Evolving Metadata Architecture

Numerous MISB metadata standards address the needs of the ISR Motion Imagery community. New applications demand additions to these standards as well as the creation of new ones. This growth challenges the current processes whereby incremental additions of new metadata items and domains occur over time where needed. This produces less than ideal utility and reusability of domains; moreover, their combination does not address the needs in complex applications. To illustrate, components comprising a system may use different temporal time-bases for their measurements. Having a metadata architecture which can support more than one timing system when associating various times with different measurements is both beneficial and essential.

To support this evolution with a more structured and organized approach, MISP-2020.1 introduced the Motion Imagery Metadata model. The MIMD model defines a modelling-base methodology to metadata, which holistically coordinates domain metadata providing the infrastructure to support new application demands. As a comprehensive metadata structure, the MIMD model defines a single model to support all existing domains and collections as well as future extensions. In use, collections define specific profiles for their applications by requiring certain model sub-structures. The MIMD model does not replace existing standards but adds an alternative for metadata collections discussed in Section 4.4.3. Expectations are that over time MIMD will become a preferred approach in developing more complex systems.

The MIMD model is in active development. Many elements of the model reflected in the standards supporting it are complete and approved for use. Use cases of the model are under evaluation with tools to support the model in development. As the MIMD model matures the expectation is the model will become the preferred approach to implementing metadata.

The following standards complete the initial suite of MIMD Model documents.

MISB ST 1901 [76], MISB ST 1902 [77], MISB ST 1903 [78], MISB ST 1904 [79], MISB ST 1905 [80], MISB ST 1906 [81], MISB ST 1907 [82]. MISB ST 1908 [83].

4.4.6 Metadata Registries

MISP metadata sets use Key Length Value encoding. The definition and its data type of a key is found in either the SMPTE RP 210 [84] or MISB ST 0807 [85] Metadata Registry. MISB ST 0607 [86] documents the process for the establishment and administration of the MISB Metadata Registry. Table 8 lists these documents with their corresponding references.

Table 8: Metadata Registry Documents

Document	Description
SMPTE RP 210 [84]	Metadata Element Dictionary Structure
MISB ST 0607 [86]	MISB Metadata Registry and Processes
MISB ST 0807 [85]	MISB KLV Metadata Registry

5 DISSEMINATION

Sensors produce Motion Imagery Data at distances from those who control and/or exploit it. Dissemination is the action of transmitting Motion Imagery Data from a source (i.e., imager, platform or control station) to one or more users. Transmitting Motion Imagery Data can affect end users in two ways: quality and latency.

When disseminating Motion Imagery, the compression applied to the Motion Imagery and data losses during transmission impact its quality. Similarly, data loss impacts reception of metadata. Latency is a measure of amount of the time it takes to move data from one point to another in a Motion Imagery System. Total latency is the elapsed time from an occurrence in the scene to viewing that occurrence at its destination. When total latency is significant, a platform controller may not be able to control accurately the imager(s), and an end user may not be able to coordinate with other users or Intel sources in real time. Therefore, minimizing total latency is an overarching design goal, especially for real time applications.

Although many technologies exist for transmitting Motion Imagery, MISB recommends the industry standard Internet Protocols (IP) – specifically, the User Datagram Protocol (UDP [87]), and the Transmission Control Protocol (TCP) [88].

5.1 Internet Protocols

The Internet Protocols constitute a family of protocols within an Internet packet-switching network to transmit data from one system to another. Table 9 provides information about the Internet Protocol family.

Table 9: Internet Protocols

Protocol Name	Description
Internet Protocol (IP)	The principal communications protocol for relaying packets across Internet networks. IP data packets (datagrams) travel from a transmitting to a receiving system using switches and routers. IP is a low-level protocol not used directly by applications; however, IP enables other higher-level protocols, such as UDP and TCP. When IP is used to send a series of datagrams, there is no guarantee that the datagrams will be delivered, or they will be received in the same order they were sent, or that when they are received the datagrams will be correct (i.e., datagrams could have been undetectably corrupted).
User Datagram Protocol (UDP/IP)	UDP is a simple transport layer protocol based on IP for delivering UDP packets. UDP packets add limited error checking and port information to IP packets. UDP does not guarantee data delivery, or that data packets arrive in order, or that packets arrive with correct data. UDP enables one-to-one and one-to-many communication. Multicast is data sent from one system to multiple systems. UDP provides one of the lowest latency methods of transmitting data to a receiver, which makes it suitable for time-sensitive and streaming data applications. Delivering Motion Imagery Data to multiple systems at once using multicast reduces overall network bandwidth.

Ordered Data Transport Protocol	In simplex communications, this protocol type offers low latency with packet order information, to facilitate minimizing network packet jitter provided by encapsulating UDP within RTP. Optional Forward Error Correction (FEC) improves reliability at the expense of an increase in latency.
Reliable Data Transport Protocol	In duplex communications, this protocol type builds on Ordered Data Transport by adding retransmission, thus further improving reliability.
Transmission Control Protocol (TCP/IP)	TCP is a transport layer protocol that provides reliable guaranteed delivery of data, with packets in order and with error checking by validating every received packet with the data sender; TCP does incur latency. Therefore, TCP does not guarantee time-sensitive delivery of data but finds use in the transfer of non-time-sensitive data, such as Motion Imagery Data files.

The method for disseminating Motion Imagery Data depends on the application. Where real time performance is critical it may be necessary to use UDP/IP. Although UDP/IP is an unreliable form of communication, there are numerous technologies and approaches to limit the impact of data errors introduced into UDP/IP (see Section 5.3.3 and the Motion Imagery Handbook). TCP/IP is an appropriate method for latency-tolerant and error-free application.

5.1.1 Real Time Protocol (RTP)

Real Time Protocol (RTP) [89] addresses delivery of A/V (audio/video) services over UDP/IP typically in bandwidth-constrained environments. In using RTP, delivery of each media type (i.e., Motion Imagery, metadata, and audio) is an independent RTP/UDP/IP data stream. A companion protocol called Real Time Control Protocol (RTCP) publishes relational timing information to synchronize the individual data streams at a receiver. Another RTP use case is encapsulating MPEG-2 TS in RTP to improve transmission robustness, where its timestamp and packet number aide packet loss and out-of-order packet detection.

5.2 Transmitting Class 0 Motion Imagery

The transmission of Class 0 Motion Imagery generally occurs between components within a Motion Imagery System, such as imager to encoder, decoder to display, etc., or within a production facility supported by a high-bandwidth wired infrastructure. For inter-equipment connection, Serial Digital Interface (SDI), GigE Vision [26], UDP/IP, and a variety of other technologies are options. RTP/UDP/IP transmission is common in long haul (many miles) transmission. MISB ST 0605 [25] provides guidance for carriage of Motion Imagery, audio, and metadata over SDI. MISB ST 1608 [27] provides guidance for carriage of Motion Imagery and metadata over GigE Vision.

5.3 Transmitting Class 1 Motion Imagery

The dissemination of Class 1 Motion Imagery includes both wired and wireless transmission. In applications requiring real-time delivery, dissemination of Class 1 Motion Imagery is MPEG-2 Transport Stream (MPEG-2 TS) over UDP/IP, MPEG-2 Transport Stream over RTP/UDP/IP, or RTP with RTCP. Non-time-sensitive applications may use TCP/IP for delivery, such as Adaptive Bitrate Streaming (ABR), which supports reliable delivery at the expense of latency.

5.3.1 MPEG-2 Transport Stream over UDP/IP

The MPEG-2 Transport Stream (MPEG-2 TS) is a widely used container for disseminating Motion Imagery Data. For example, Motion Imagery Data transmitted from an airborne platform is typically in a MPEG-2 TS container, as well as to points within a network that support exploitation. MPEG-2 TS is a common container for delivering Motion Imagery Data over IP as well. The MISP provides guidance on inserting MPEG-2 TS packets into UDP packets, see MISB ST 1402 [44]. This guidance assures off-the-shelf products can properly decode the Motion Imagery Data.

5.3.2 MPEG-2 Transport Stream over RTP/UDP/IP

Encapsulating MPEG-2 TS in RTP takes advantage of the timestamp and packet number contained in the RTP header. MISB ST 0804 [90], provides information for encapsulating Motion Imagery Data and MPEG-2 TS in RTP and is the basis for the following requirement.

Requirement	
MISP-2015.1-76	A MPEG-2 Transport Stream encapsulated in Real Time Protocol (RTP) shall comply with MISB ST 0804.

5.3.2.1 RTP with RTCP

RTP with RTCP provides a second means to deliver Class 1 Motion Imagery. MISB ST 0804 [90] provides guidance on the proper use of RTP for the delivery of Motion Imagery, metadata, and audio.

Requirement(s)	
MISP-2015.1-77	Class 1 Motion Imagery encapsulated in Real Time Protocol (RTP) shall comply with MISB ST 0804.
MISP-2015.1-78	Class 2 Motion Imagery encapsulated in Real Time Protocol (RTP) shall comply with MISB ST 0804.

MISB RP 1302 [91] describes a method to insert KLV-encoded metadata in a Session Description Protocol (SDP), and is recommended guidance when adding security markings to an RTP stream.

5.3.2.2 Error Handling

Monitoring the health of a delivery system is important in detecting irregularities in received data. Both MPEG-2 TS and RTP provide mechanisms to aid detection of missing and lost data packets. How a decoder processes either missing or damaged packets depends on its implementation. Many MISP defined KLV metadata sets include error detection (i.e., cyclic redundancy check or CRC) to validate the transmitted and received data match. In detecting errors, it is important receivers properly handle the error and disregard erroneous data.

5.3.3 Retransmission Protocols

Introducing selective packet retransmission (reliable UDP) to recover from packet loss is one method to ensure delivery of data over IP. Automatic Repeat Request (ARQ) protocols allow systems to detect data stream discontinuities and retransmit lost data within a short window. Two

ARQ protocols -- the web-based Quick UDP Internet Connections (QUIC) [92] and generic packet-based Secure Reliable Transport (SRT) [93] -- are on the standards track within the Internet Engineering Task Force (IETF). Generally, such methods introduce latency into the stream, thus limiting their practicality in time-critical applications. Trading latency for improved Motion Imagery reception where possible, however, significantly improves received data reliability.

5.3.3.1 Quick UDP Internet Connections (QUIC)

QUIC aims to be nearly equivalent to a TCP connection but with much-reduced latency. Two goals of QUIC include: 1) reduce the overhead during connection setup, and 2) using UDP as its basis each QUIC stream is flow controlled and data retransmitted at the level of QUIC, not UDP.

5.3.3.2 Secure Reliable Transport (SRT)

SRT improves point-to-point and point-to-multipoint transmission of video, and load-balancing across multiple links with seamless switching. SRT also protects transmissions from eavesdropping using encryption between sender and receiver. Forward Error Correction (FEC) in SRT ensures further data reliability.

5.3.4 Adaptive Bitrate Streaming (ABR)

ABR is a media-streaming model for delivery of media content in which control lies with the client. MISB ST 1910 describes the encoding of H.265/H.264 compressed imagery, audio, and KLV metadata within a Common Media Application Format (CMAF) file for ABR delivery. ABR is a TCP/IP technology intended for non-time critical applications. Delivery of CMAF files is application specific utilizing technologies such as Dynamic Adaptive Streaming over HTTP (MPEG-DASH) and HTTP Live Streaming (HLS [94]); ST 1910 defines the packaging of content for delivery using these methods.

Requirement	
MISP-2020.3-120	Class 1 Motion Imagery, audio, and metadata encapsulated per CMAF ISO/IEC 23000-19: 2020 for adaptive bitrate streaming shall comply with MISB ST 1910.

5.4 Transmitting Class 2 Motion Imagery

One method for delivery of Class 2 Motion Imagery is JPEG2000 Interactive Protocol (JPIP), which relies on TCP/IP. For H.265/H.264 compressed imagery the various methods appropriate for Class 1 Motion Imagery (see Section 5.3) apply to Class 2 Motion Imagery as well. An option is to convert Class 2 Motion Imagery to Class 1 Motion Imagery as described in Section 6.3. Finally, GigE Vision is a solution for Class 2 Motion Imagery localized transmission, such as a sensor to a processor sub-system.

5.4.1 JPEG2000 Interactive Protocol (JPIP)

JPEG2000 Interactive Protocol (JPIP) is a delivery option for Class 2 Motion Imagery. JPIP applies in systems that collect and compress Class 2 Motion Imagery using JPEG 2000 with dissemination over reduced bandwidth links. MISB RP 0811 [95] defines a JPEG 2000 Interactive Protocol (JPIP) Profile for client/server interaction. It defines the expected behavior for client/server interactions for JPEG 2000 compressed imagery within the context of the JPIP protocol. RP 0811 does not address the delivery of metadata items within JPIP.

Requirement	
MISP-2015.1-79	JPEG2000 Interactive Protocol (JPIP) shall be implemented in accordance with MISB RP 0811.

Class 2 Motion Imagery converted into Class 1 Motion Imagery may use the delivery methods described in Section 5.3.

5.5 Transmitting Standalone Metadata

MISB ST 0804 [90] provides guidance for transmitting metadata over RTP independent of Motion Imagery. Receivers of KLV can detect if errors have occurred. MISB ST 1608 [27] provides guidance for transmitting metadata using GigE Vision. Other formats of metadata that do not provide error detection should use a reliable transport (i.e., TCP/IP) when possible.

6 APPLICATION SPECIFIC SYSTEMS

Systems outside of the Motion Imagery ecosphere need to communicate with Motion Imagery Data. This section identifies MISB guidance for interfacing to other application specific systems.

6.1 Cursor on Target (CoT)

Cursor on Target (CoT) is a simple messaging format for situational awareness and command-and-control functions. To facilitate interoperability, recommended conversions from ST 0601 [54] KLV metadata to two basic CoT schema messages (platform position and sensor point-of-interest (SPOI)) are presented in MISB ST 0805 [96].

Requirement	
MISP-2015.1-80	KLV to Cursor-on-Target (CoT) encoding shall be in accordance with MISB ST 0805.

6.2 STANAG 4559 Data Model

MISB RP 0813 [97] describes the necessary conditions for integration of Motion Imagery products into the STANAG 4559 [98] Data Model and Interface, which is based on the Multi-sensor Aerospace-ground Joint Intelligence, Surveillance and Reconnaissance Interoperability Coalition (MAJIIC) Coalition Shared Data (CSD).

6.3 STANAG 4586 Control of Motion Imagery Payloads

MISB ST 1101 [99] provides direction on compression, container, and delivery protocol choices for disseminating Class 1 Motion Imagery from a Class 0 Motion Imagery or Class 2 Motion Imagery source. ST 1101 references several STANAG 4586 [100] control messages for directing a Motion Imagery Sensor, selecting the compression, container, and delivery format of Class 1 Motion Imagery, and choosing metadata per mission needs.

Requirement	
MISP-2015.1-81	When requesting Class 1 Motion Imagery or Class 2 Motion Imagery from a platform that supports MISB ST 1101, the messages shall comply with MISB ST 1101.

6.4 AAF Profile for Aerial Surveillance and Photogrammetry Applications (ASPA)

The AAF Profile for Aerial Surveillance and Photogrammetry Applications (ASPA) leverages MXF and forms the basis of the AAF format in the NGA Image Product Library (IPL). MISB ST 0301 [101] provides guidance and rules for producing ASPA files.

As the MXF standards afford many options to support many application domains, the following requirements – per SMPTE ST 378 [102], SMPTE 391 [103], SMPTE ST 379-1 [104], SMPTE ST 381-1 [105], and MISB ST 0301 – limit MXF to a minimum level to achieve interoperability between the implementing entities.

Requirement(s)	
MISP-2015.1-12	For file exchange, operational patterns 1a (OP-1a) and 1b (OP-1b) as per SMPTE ST 378 and SMPTE ST 391, respectively, shall be used.
MISP-2015.1-13	Motion Imagery, metadata and audio shall use the method of frame-based mapping within the generic container in accordance with SMPTE ST 379-1 and SMPTE ST 381-1.
MISP-2015.1-14	All data constraints for an ASPA MXF file shall comply with MISB ST 0301.
MISP-2015.1-15	Although possibly not interpreted, an MXF player shall accept dark (unknown) data.

Appendix A Deprecated/Retired Standards & Requirements

A.1 Deprecated/Retired Standards Documents

The standards documents listed in this section are legacy. Legacy indicates outdated technologies no longer recommended for new implementations; such standards documents have a status of “Deprecated.” A “Deprecated” standard remains in standing for a period to maintain backward compatibility. Assuming a 5-10-year technology refresh cycle, at that time a “Deprecated” standard will move to the status of “Retired.” New implementations should not use a “Retired” standard after it expires.

A standard may be either superseded by a new revision of that standard or replaced by a new standard. A superseded standard is “Deprecated.” A standard recast in the form of one or more requirements no longer exist as standalone references and are likewise “Deprecated.”

Document	Title	Status	Effective Date
MISB EG 0104	Predator UAV Basic Universal Metadata Set	Deprecated	Sep 2008
MISB RP 0103	Timing Reconciliation Universal Metadata Set for Digital Motion Imagery	Deprecated	May 2009
MISB RP 0608	Motion Imagery Identification [Superseded by MISB ST 1204]	Deprecated	Oct 2013
MISB RP 0705	LVSD Compression Profile [Superseded by BPJ2K01.10]	Deprecated	Oct 2013
MISB EG 0810	Profile 2: KLV for LVSD Applications	Deprecated	Oct 2013
MISB ST 9701	MPEG-2 Transport Stream [Superseded by MISB ST 1402]	Deprecated	Feb 2014
MISB ST 9708	Imbedded Time Reference for Motion Imagery Systems [Superseded by MISB ST 0604]	Deprecated	Feb 2014
MISB ST 9715	Time Reference Synchronization [Superseded by MISB ST 0603]	Deprecated	Feb 2014
MISB RP 9717	Packing KLV Packets into MPEG-2 Systems Streams [Superseded by MISB ST 1402]	Deprecated	Feb 2014
MISB RP 0101	Use of MPEG-2 System Streams in Digital Motion Imagery Systems [Superseded by MISB ST 1402]	Deprecated	Feb 2014
MISB ST 9601	Standard Definition Digital Motion Imagery, Compression Systems [Recast as REQs MISP-2015.1-32, -33, -34, -35, -47, -50]	Deprecated	Oct 2014
MISB ST 9702	Standard Definition Digital Motion Imagery Sampling Structure [Recast as REQ MISP-2015.1-18]	Deprecated	Oct 2014
MISB ST 9703	Digital Motion Imagery, Uncompressed Baseband Signal Transport and Processing [Recast as REQs MISP-2015.1-23, -25, -31]	Deprecated	Oct 2014

MISB ST 9704	Digital Motion Imagery, Compression Conversions [Recast as REQ MISP-2015.1-17]	Deprecated	Oct 2014
MISB ST 9705	Standard Definition Digital Motion Imagery, Format Conversions [Recast as REQ MISP-2015.1-17]	Deprecated	Oct 2014
MISB ST 9706	Motion Imagery Still Frames [Recast as REQ MISP-2015.1-16]	Deprecated	Oct 2014
MISB ST 9707	Standard Definition Digital Motion Imagery Tape Recorder, Digital Motion Imagery Servers, and Similar Systems Input / Output Protocol [Recast as REQ MISP-2015.1-23]	Deprecated	Oct 2014
MISB ST 9709	Use of Closed Captioning for Core Metadata Analog Video Encoding [Recast as REQ MISP-2015.1-83 (Deprecated)]	Deprecated	Oct 2014
MISB ST 9710	High Definition Television Systems (HDTV) [Recast as REQs MISP-2015.1-02,-20,-21,-25]	Deprecated	Oct 2014
MISB ST 9711	Intelligence Motion Imagery Index, Geospatial Metadata	Deprecated	Oct 2014
MISB ST 9712	Intelligence Motion Imagery Index, Content Description Metadata (Dynamic Metadata Dictionary Structure and Contents) [Superseded by MISB ST 0607 & ST 0807]	Deprecated	Oct 2014
MISB ST 9713	Data Encoding Using Key-Length-Value [Recast as REQ MISP-2015.1-07]	Deprecated	Oct 2014
MISB ST 9714	Time Code Embedding	Deprecated	Oct 2014
MISB ST 9716	Packing KLV Packets into SMPTE 291 Ancillary Data Packets [Superseded by MISB ST 0605]	Deprecated	Oct 2014
MISB ST 9718	Packing KLV Packets into AES3 Serial Digital Audio Streams	Deprecated	Oct 2014
MISB ST 9719	Analog Video Migration [Recast as REQs MISP-2015.1-06, -17, -67]	Deprecated	Oct 2014
MISB RP 9720	Motion Imagery System Matrix (MISM)	Deprecated	Oct 2014
MISB RP 9721	Motion Imagery Tape Formats	Deprecated	Oct 2014
MISB ST 9723	Compressed High Definition Advanced Television (ATV) and associated Motion Imagery Systems [Recast as REQs MISP- 2015.1-32,-33,-34,-35,-36,-38,-39,-40,-41,-47]	Deprecated	Oct 2014
MISB ST 9803	Serial Data Transport Interface	Deprecated	Oct 2014
MISB ST 9811	Progressively Scanned Enhanced Definitio1002n Digital Motion Imagery [Recast as REQ MISP-2015.1-19]	Deprecated	Oct 2014

MISB ST 9901	Fiber Optic Interfaces Uncompressed Baseband Signal Transport and Processing [Recast as REQ MISP-2015.1-84 (Deprecated)]	Deprecated	Oct 2014
MISB RP 9902	Authorized Limited Applications of DV Format Video	Deprecated	Oct 2014
MISB RP 0106	Advanced Authoring Format [Recast as REQ MISP-2015.1: U.S. Specific-01	Deprecated	Oct 2014
MISB RP 0108	Material Exchange Format [Recast as REQs MISP-2015.1-12,-13,-14,-15]	Deprecated	Oct 2014
MISB ST 0201	Uncompressed Enhanced Motion Imagery Baseband Signal Transport [Recast as REQ MISP-2015.1-24]	Deprecated	Oct 2014
MISB ST 0202	Compressed Enhanced Definition Advanced Television (ATV) and associated Motion Imagery Systems [Recast as REQs MISP-2015.1-32,-33,-34,-35,-38,-39,-40,-41,-47]	Deprecated	Oct 2014
MISB RP 0401	Infrared Motion Imagery System Matrix (IR MISM)	Deprecated	Oct 2014
MISB RP 0402	Parallel Interface for Infrared Motion Imagery	Deprecated	Oct 2014
MISB RP 0606	Authorized Use of JPEG 2000 for Large Volume Streaming Data Imagery [Recast as REQ MISP-2015.1-51]	Deprecated	Oct 2014
MISB RP 1004	LVSD Motion Imagery System Matrix (LVSD MISM)	Deprecated	Oct 2014
MISB RP 1203	Video Interpretability and Quality Measurement and Prediction	Deprecated	Oct 2019
MISB RP 0701	Common Metadata System: Structure	Deprecated	Oct 2020

A.2 Deprecated Requirements

This section lists requirements found across all versions of the MISP which are no longer in force.

Requirement(s)	
MISP-2015.1-82 (Deprecated)	One AES3 audio channel (one stereo pair) encoded in accordance with SMPTE ST 291 and carried within the Horizontal Ancillary Space (HANC) of a Serial Digital Interface (SDI) shall be reserved for metadata.
MISP-2015.1-83 (Deprecated)	Encoding metadata on Line 21 of the vertical interval of legacy system video shall be in accordance with CEA-608 and CEA-708.
MISP-2015.1-84 (Deprecated)	The transport of SMPTE ST 259, SMPTE ST 292-1, and SMPTE ST 424 signals over fiber optic shall comply with SMPTE ST 297.
MISP-2015.1-40 (Deprecated)	A Class 1 Motion Imagery Decoder shall support the decoding of Class 1 Motion Imagery compressed using H.264/AVC Baseline Profile, Main Profile and High Profile up to and including Level 4.1.
MISP-2015.1-33 (Deprecated)	Class 1 Motion Imagery compressed using H.264/AVC shall use Baseline, Main or High Profile.

MISP-2015.3-85 (Deprecated)	A Class 1 Motion Imagery Decoder shall support the decoding of Class 1 Motion Imagery compressed using H.264/AVC Baseline Profile, Main Profile and High Profile up to and including Level 4.
MISP-2015.1-03 (Deprecated)	All Motion Imagery shall include a Precision Time Stamp in accordance with MISB ST 0603.
MISP-2015.1-04 (Deprecated)	Any metadata that contains a timestamp item shall include a Precision Time Stamp in accordance with MISB ST 0603.
MISP-2015.1-43 (Deprecated)	Class 1 Motion Imagery compressed using H.264/AVC shall contain a Precision Time Stamp in accordance with MISB ST 0604.
MISP-2015.1-44 (Deprecated)	Class 1 Motion Imagery compressed using MPEG-2 shall contain a Precision Time Stamp in accordance with MISB ST 0604.
MISP-2016.1-86 (Deprecated)	Class 2 Motion Imagery compressed using H.264/AVC shall contain a Precision Time Stamp in accordance with MISB ST 0604.
MISP-2016.1-87 (Deprecated)	Class 2 Motion Imagery compressed using H.262/MPEG-2 shall contain a Precision Time Stamp in accordance with MISB ST 0604.
MISP-2015.1-58 (Deprecated)	Class 3 Motion Imagery shall be compressed using H.264/AVC or H.262/MPEG-2.
MISP-2015.1-59 (Deprecated)	When timestamp information is available in Class 3 Motion Imagery, it shall be converted to a Precision Time Stamp in accordance with MISB ST 0603.
MISP-2015.1-60 (Deprecated)	When a Precision Time Stamp is generated from Class 3 Motion Imagery, it shall be inserted into a H.264/AVC or H.262/MPEG-2 compressed elementary stream in accordance with MISB ST 0604.
MISP-2015.1-61 (Deprecated)	When timestamp information is available in Class 3 Motion Imagery and after conversion to a Precision Time Stamp in accordance with MISB ST 0603, it shall be inserted as metadata in accordance with MISB ST 0601.
MISP-2018.1-100 (Deprecated)	While compressing Class 1 Motion Imagery with H.265/HEVC, the compression profile shall be Main 10, which includes the profile Main.
MISP-2017.1-93 (Deprecated)	While compressing Class 1 Motion Imagery with H.264/AVC, the compression profile shall be Constrained Baseline, Main or High Profile.
MISP-2018.1-101 (Deprecated)	At least one Point of Interoperability (POI) (H.265/HEVC Table 6) shall be supported by a Class 1 Motion Imagery [H.265/HEVC] Encoder.
MISP-2015.1-35 (Deprecated)	While compressing Class 1 Motion Imagery with H.262/MPEG-2, the compression profile shall be Main or High Profile.
MISP-2015.1-36 (Deprecated)	At least one Point of Interoperability (POI) (H.264/AVC Table 7) shall be supported by a Class 1 Motion Imagery [H.264/AVC] Encoder.
MISP-2015.1-23 (Deprecated)	The Container format for Standard Definition (SD) Class 0 Motion Imagery for source aspect ratios of 4:3 and 16:9 shall be defined by SMPTE ST 259 Levels C and D respectively.
MISP-2015.1-24 (Deprecated)	The Container format for Enhanced Definition (ED) 525- and 625-line Class 0 Motion Imagery over SMPTE ST 292-1 shall be defined by SMPTE ST 349.
MISP-2015.1-25 (Deprecated)	The Container format for High Definition (HD) Class 0 Motion Imagery up to 1080p 60 frames per second shall be defined by SMPTE ST 424. The Container format for High Definition Class 0 Motion Imagery up to 1080p at 30 frames per second may use SMPTE ST 292-1.

MISP-2015.1-26 (Deprecated)	The Container format for High Definition (HD) Class 0 Motion Imagery up to 1080p at 60 frames per second shall be defined by SMPTE ST 424 with signal mapping specified in SMPTE ST 425-1.
MISP-2015.1-27 (Deprecated)	The Container format for Ultra High Definition (UHD) Class 0 Motion Imagery shall be defined by SMPTE ST 435-2 with signal mapping specified in SMPTE ST 2036-3.
MISP-2015.1-28 (Deprecated)	The mapping of digital infrared Class 0 Motion Imagery into a SMPTE ST 292-1 interface shall be defined by MISB ST 0403. Note: Requirement moved to MISB ST 0605.9
MISP-2015.1-29 (Deprecated)	Class 0 Motion Imagery carried by a SMPTE ST 292-1 or SMPTE ST 424 Container shall contain a Precision Time Stamp in accordance with MISB ST 0605.
MISP-2015.1-30 (Deprecated)	Class 0 Motion Imagery metadata inserted into SMPTE ST 292-1 or SMPTE ST 424 shall comply with MISB ST 0605.
MISP-2015.1-31 (Deprecated)	One AES3 audio channel (one stereo pair) encoded in accordance with SMPTE ST 291 and inserted into the Horizontal Ancillary Space (HANC) of a Serial Digital Interface (SDI) shall be reserved for mission audio (narration, etc.). Requirement moved to ST 0605.9.
MISP-2015.1-52 (Deprecated)	Class 2 Motion Imagery compressed with H.264/AVC shall use High 4:2:2 Profile (Hi422P) or High 4:4:4 Predictive Profile (Hi444PP).
MISP-2018.1-96 (Deprecated)	Motion Imagery shall include a timestamp representing Absolute Time consistent with MISB ST 0603.
MISP-2018.1-105 (Deprecated)	Class 2 Motion Imagery compressed with H.265/HEVC shall use the profile Main 4:2:2 12 or Main 4:4:4 12.
MISP-2018.1-106 (Deprecated)	A Class 2 Motion Imagery Decoder for H.265/HEVC shall support the decoding of the profiles Main 4:2:2 12 and Main 4:4:4 12 up to and including Level 5.1.
MISP-2015.1-17 (Deprecated)	Analog Motion Imagery that conforms to NTSC [8] or PAL [9] formats and is converted to digital shall comply with ITU-R BT.601 component (4:2:2).
MISP-2015.1-18 (Deprecated)	The Motion Imagery sampling format for Standard Definition (SD) Motion Imagery shall be defined by ITU-R BT.601.

Appendix B References

MISB references cited here-in reflect versions current to the publication date of a document. In the event of a MISB document correction, the corrected document will have a single letter Minor Version appended to the Major Version number per the MISB Document Development Process. For example, corrections to ST 0001.2, which has a Major Version of 2, becomes ST 0001.2a, which includes a Minor Version of “a”. The MISB will not update the referring document (this document) with the Minor Version number change. When acquiring any MISB reference listed below from an NGA repository the latest version may be either a Major or Minor Version.

[1] NATO STANAG 4609: NATO Digital Motion Imagery Standard, Edition 5, 30 Jul 2020.

[2] MISB MISP-2021.1: U.S. Governance, Oct 2020.

[3] MISB MISP-2021.1: U.S. Specific, Oct 2020.

[4] MISB MISP-2021.1: Motion Imagery Handbook, Oct 2020.

[5] SMPTE EG 28-1993 Annotated Glossary of Esential Terms for Electronic Production.

[6] MISB TRM 1404.1 H.264 Compression Principles, Jun 2015.

[7] MISB RP 0904.4 Bandwidth/Quality/Latency Tradeoffs for Compressed Motion Imagery, Nov 2018.

[8] SMPTE ST 170:2004 Television - Composite Analog Video Signal - NTSC for Studio Applications.

[9] ITU-R BT.470-7 (02/2005) Conventional Analogue Television Systems.

[10] SMPTE ST 274:2008 Television - 1920 x 1080 Image Sample Structure, Digital Representation and Digital Timing Reference Sequence for Multiple Picture Rates.

[11] MISB ST 1507.2 Motion Imagery Sensor Timing Metadata, Oct 2017.

[12] MISB ST 0603.5 MISP Time System and Timestamps, Oct 2017.

[13] SMPTE ST 336:2017 Data Encoding Protocol Using Key-Length-Value.

[14] MISB ST 0107.4 KLV Metadata in Motion Imagery, Feb 2019.

[15] MISB ST 1201.4 Floating Point to Integer Mapping, Feb 2019.

[16] MISB ST 1303.2 Multi-Dimensional Array Pack, Jun 2020.

[17] MIL-STD-2500C (CN2) National Imagery Transmission Format Version 2.1 for the National Imagery Transmission Format, Change Notice 2, 02 Jan 2019.

[18] NATO STANAG 4545 NATO Secondary Imagery Format (NSIF), Ed. 2, May 2013.

[19] MISB RP 1011.1 LVSD Motion Imagery Streaming, Feb 2014.

[20] ITU-R BT.1358-1 (09/2007) Studio Parameters of 625 and 525 Line Progressive Television.

[21] SMPTE ST 296:2012 1280 x 720 Progressive Image 4:2:2 and 4:4:4 Sample Structure - Analog and Digital Representation and Analog Interface.

[22] SMPTE ST 2036-1:2014 Ultra High Definition Television - Image Parameter Values for Program Production.

[23] MISB ST 0404.1 Compression for Infrared Motion Imagery, Feb 2014.

[24] SMPTE ST 291-1:2011 Ancillary Data Packet and Space Formatting.

[25] MISB ST 0605.10 Class 0 Motion Imagery Metadata and Audio over SDI, Oct 2020.

[26] Advanced Vision+Imaging (AIA) GigE Vision Specification 2.1 Video Streaming and Device Control Over Ethernet Standard, Nov 2018.

[27] MISB ST 1608.1 Transport of Motion Imagery and Metadata over GigE Vision, Oct 2017.

[28] SMPTE ST 377-1:2011 Material Exchange Format (MXF) - File Format Specification.

[29] MISB ST 1606.1 MXF Profile for High Performance Motion Imagery Applications, Oct 2017.

[30] ISO/IEC 23008-2:2017 Information Technology - High efficiency coding and media delivery in heterogeneous environments - Part 2: High efficiency video coding.

[31] ITU-T H.265 (V7) (11/2019) High efficiency video coding.

[32] ISO/IEC 14496-10:2014 Information Technology - Coding of audio-visual objects - Part 10: Advanced Video Coding.

[33] ITU-T H.264 (V13) (06/2019) Advanced video coding for generic audiovisual services.

[34] ISO/IEC 13818-2:2013 Information Technology - Generic coding of moving pictures and associated audio information: Video.

[35] ITU-T H.262 (V3) (02/2012) Generic coding of moving pictures and associated audio information: Video.

[36] MISB RP 0802.2 H.264/AVC Motion Imagery Coding, Feb 2014.

[37] MISB ST 0602.4 Annotation Metadata Set, Feb 2014.

[38] MISB ST 0604.6 Timestamps for Class 1/Class 2 Motion Imagery, Oct 2017.

[39] MISB ST 1001.1 Audio Encoding, Feb 2014.

[40] ISO/IEC 13818-1:2019 Information technology - Generic coding of moving pictures and associated audio information: Systems.

[41] ITU-T H.222.0 (V7) (08/2018) Generic Coding of Moving Pictures and Associated Audio Information: Systems.

[42] ISO/IEC 13818-4:2004 Information Technology - Generic coding of moving pictures and associated audio information: Compliance Testing.

[43] MISB TRM 0909.4 Constructing a MISP Compliant File/Stream, Oct 2014.

[44] MISB ST 1402.2 MPEG-2 Transport Stream for Class 1/Class 2 Motion Imagery, Audio and Metadata, Oct 2016.

[45] ISO/IEC 23000-19:2020 Information technology - Multimedia application format (MPEG-A) - Part 19: Common media application format (CMAF) for segmented media.

[46] ISO/IEC 23009-1 :2019 Information technology - Dynamic adaptive streaming over HTTP (DASH) - Part 1: Media presentation description and segment formats.

[47] MISB ST 1910.1 Adaptive Bitrate Content Encoding, Oct 2020.

[48] ISO/IEC BIIF Profile BPJ2K.01.10 BIIF Profile for JPEG 2000, Apr 2009.

[49] ISO/IEC 15444-4:2004 JPEG2000 Image Coding System - Part 4: Conformance Testing.

[50] NGA.STND.0044_1.3 Motion Imagery Extension for NITF 2.1 Reference, Feb 2019.

[51] MISB ST 1909.1 Metadata Overlay for Visualization, Oct 2020.

[52] MISB ST 1204.3 Motion Imagery Identification System (MIIS) Core Identifier, Feb 2020.
[53] MISB ST 1301.2 Motion Imagery Identification System - Augmentation Identifiers Local Set, Feb 2014.
[54] MISB ST 0601.17 UAS Datalink Local Set, Oct 2020.
[55] MISB ST 0903.5 Video Moving Target Indicator and Track Metadata, Feb 2020.
[56] MISB ST 0808.2 Ancillary Text Metadata Sets, Oct 2015.
[57] MISB ST 0806.4 Remote Video Terminal Metadata Set, Feb 2014.
[58] MISB ST 0801.8 Photogrammetry Metadata Set for Motion Imagery, Oct 2019.
[59] MISB ST 1107.3 Metric Geopositioning Metadata Set, Feb 2018.
[60] MISB ST 1202.2 Generalized Transformation Parameters, Feb 2015.
[61] MISB ST 1010.3 Generalized Standard Deviation and Correlation Coefficient Metadata, Oct 2016.
[62] MISB ST 1002.2 Range Motion Imagery, Jun 2016.
[63] MISB ST 0809.2 Meteorological Metadata Local Set, Nov 2018.
[64] MISB ST 1108.3 Quality and Interpretability Local Set, Oct 2019.
[65] MISB ST 0901.3 Video-National Imagery Interpretability Rating Scale, Oct 2017.
[66] MISB ST 0102.12 Security Metadata Universal and Local Sets for Motion Imagery Data, Jun 2017.
[67] MISB ST 1206.1 SAR Motion Imagery Metadata, Jun 2017.
[68] MISB ST 1403.2 SARMI Threshold Metadata Sets, Nov 2018.
[69] MISB ST 1603.2 Time Transfer Pack, Oct 2017.
[70] MISB ST 1504.1 Natural Representation of Orbital State Vectors, Nov 2018.
[71] MISB ST 0902.8 Motion Imagery Sensor Minimum Metadata Set, Nov 2018.
[72] MISB ST 1801.1 Motion Imagery Metadata (MIMD) Model: Surface Profile, Oct 2020.
[73] MISB ST 1607 Constructs to Amend/Segment KLV Metadata, Oct 2016.
[74] MISB ST 1601.1 Geo-Registration Local Set, Nov 2018.
[75] MISB ST 1602.1 Composite Imaging Local Set, Nov 2018.
[76] MISB ST 1901.1 Motion Imagery Metadata (MIMD) Modelling Rules, Jun 2020.
[77] MISB ST 1902.1 Motion Imagery Metadata Model-to-KLV Instructions, Jun 2020.
[78] MISB ST 1903.1 Motion Imagery Metadata (MIMD): Model, Jun 2020.
[79] MISB ST 1904.1 Motion Imagery Metadata (MIMD): Base Attributes, Jun 2020.
[80] MISB ST 1905.1 Motion Imagery Metadata (MIMD) Model: Platforms, Jun 2020.
[81] MISB ST 1906.1 Motion Imagery Metadata (MIMD): Staging System, Jun 2020.
[82] MISB ST 1907.1 Motion Imagery Metadata (MIMD) Model: Payloads, Jun 2020.
[83] MISB ST 1908.1 Motion Imagery Metadata (MIMD) Model: Imager System, Jun 2020.
[84] SMPTE RP 210v13:2012 Metadata Element Dictionary.
[85] MISB ST 0807.25 MISB KLV Metadata Registry, Jun 2020.
[86] MISB ST 0607.5 MISB Metadata Registry and Processes, Feb 2019.
[87] IETF RFC 768 User Datagram Protocol, Aug 1980.

[88] IETF RFC 793 Transmission Control Protocol, Sep 1981.

[89] IETF RFC 3550 RTP: A Transport Protocol for Real-Time Applications, Jul 2003.

[90] MISB ST 0804.4 Real-Time Protocol for Motion Imagery and Metadata, Feb 2014.

[91] MISB RP 1302 Inserting KLV in Session Description Protocol (SDP), Feb 2014.

[92] IETF draft-ietf-quic-transport-22 QUIC: A UDP-Based Multiplexed and Secure Transport, Jul 2019.

[93] IETF draft-sharabayko-mops-srt-00 The SRT Protocol, Mar 2020.

[94] IETF RFC 8216 HTTP Live Streaming 2nd Edition, Apr 2020.

[95] MISB RP 0811.1 JPIP Profile (Client/Server Functions), Oct 2014.

[96] MISB ST 0805.1 KLV to Cursor-on-Target (CoT) Conversions, Feb 2014.

[97] MISB RP 0813.1 Integration of Motion Imagery into the STANAG 4559 Data Model, Feb 2014.

[98] NATO STANAG 4559 NATO Standard ISR Library Interface Edition 3, Nov 2010.

[99] MISB ST 1101 STANAG 4586 Control on UAS Motion Imagery Payloads, Oct 2014.

[100] NATO STANAG 4586 Edition 3, 12 Nov 2010.

[101] MISB ST 0301.5 MISB Profile for Aerial Surveillance and Photogrammetry Applications (ASPA), Jul 2010.

[102] SMPTE ST 378:2004 Television - Material Exchange Format (MXF) - Operational Pattern 1A (Single Item, Single Package).

[103] SMPTE ST 391:2004 Television - Material Exchange Format (MXF) – Operational Pattern 1b (Single Item, Ganged Packages).

[104] SMPTE ST 379-1:2009 Material Exchange Format (MXF) - MXF Generic Container.

[105] SMPTE ST 381-1:2005 Material Exchange Format (MXF) - Mapping MPEG streams into the MXF Generic Container.

[106] MISB TRM 1605 NATO STANAG 4609 Edition 3 Mapped to NATO STANAG 4609 Edition 4, Oct 2016.

[107] MISB TRM 1803 NATO STANAG 4609 Edition 4 Mapped to NATO STANAG 4609 Edition 5, Nov 2018.

[108] MISB TRM 1802 NATO MAJIIC 2 STANAG 4609 Motion Imagery Implementation Guide, Feb 2018.

[109] MISB Document Development Process 7.0, Feb 2020.

[110] MISB Document Author Kit 5.0, Feb 2020.

Appendix C NATO Traceability and Implementation Information

C.1 Edition Mapping of MISP to STANAG 4609

C.1.1 Edition Mapping of STANAG 4609 Ed 3 to STANAG 4609 Ed 4

MISB TRM 1605 [106] defines the mapping of NATO STANAG 4609 Edition 3, and its corresponding NATO Allied Engineering Documentation Publication-8 (AEDP-8) Edition 3 standards to NATO STANAG 4609 Edition 4 standards and requirements as represented in MISP-2015.1. STANAG 4609 Edition 4 introduces a new structure of documents, where STANAG 4609 is now a cover document rather than a standalone document, which points directly to the U.S. Motion Imagery Standards Profile (MISP).

C.1.2 Edition Mapping of STANAG 4609 Ed 4 to STANAG 4609 Ed 5

MISB TRM 1803 [107] defines the mapping of standards and requirements for NATO STANAG 4609 Edition 4 as represented in MISP-2015.1 to NATO STANAG 4609 Edition 5 as represented in MISP-2019.1. As of STANAG 4609 Edition 4, the STANAG is a covering document rather than a standalone document, which points directly to a version of the U.S. Motion Imagery Standards Profile (MISP).

C.2 NATO MAJIIC 2 STANAG 4609 Motion Imagery Implementation Guide

MISB TRM 1802 [108] documents the Multi-Intelligence All-Source Joint Intelligence Surveillance and Reconnaissance Interoperability Coalition 2 (MAJIIC 2) implementation guidance for NATO STANAG 4609 as published by the NATO Communications and Information Agency (NCIA). The MAJIIC 2 project closed at the end of 2015 and this TRM serves as a repository for their artifacts.

Appendix D Acronyms

ABR	Adaptive Bitrate Streaming
ARQ	Automatic Repeat Request
ASPA	Aerial Surveillance and Photogrammetry Applications
AVC	Advanced Video Coding
BMP	Bitmap image
bps	Bits per second
CMAF	Common Media Application Format
CoT	Cursor on Target
DAR	Display Aspect Ratio
DASH	Dynamic Adaptive Streaming over HTTP
ED	Enhanced Definition
FEC	Forward Error Correction
FMV	Full Motion Video
FPS	Frames per second
Gb	Gigabit
HANC	Horizontal Ancillary Space
HD	High Definition
HEVC	High Efficiency Video Coding
HL	High Level
Hz	Hertz
IEC	International Electrotechnical Commission
IETF	Internet Engineering Task Force
IPL	NGA Image Product Library
ISO	International Standards Organization
ISR	Intelligence, Surveillance, and Reconnaissance
ITU	International Telecommunication Union
ITU-R	International Telecommunication Union - Radiocommunications
JPEG	Joint Photographic Experts Group
JPIP	JPEG 2000 Interactive Protocol
KLV	Key Length Value
LADAR	LAser Detection And Ranging SyLight Detection and Ranging
LVMI	Large Volume Motion Imagery
MAJIIC	Multi-sensor Aerospace-ground Joint Intelligence, Surveillance and Reconnaissance Interoperability Coalition
Mbps	Megabits (10^6) per second
MI	Motion Imagery
MIMD	Motion Imagery Metadata
MISB	Motion Imagery Standards Board
MISMMS	Motion Imagery Sensor Minimum Metadata Set
MISP	Motion Imagery Standards Profile
ML	Main Level
MP	Main Profile
MPEG	Moving Pictures Experts Group

MRTFB	Major Range and Test Facility Base
MXF	Material Exchange Format
NATO	North Atlantic Treaty Organization
NITF	National Imagery Transmission Format Standard
NSG	National System for Geospatial-Intelligence
NTB	NITF Technical Board
NTSC	National Television System Committee
PAL	Phase Alternating Line
PAR	Pixel Aspect Ratio
QUIC	Quick UDP Internet Connections
RP	Recommended Practice
RTCP	Real Time Control Protocol
RTP	Real Time Protocol
RVT	Remote Video Terminal
SAR	Source Aspect Ratio and Synthetic Aperture Radar
SDI	Serial Data Interface
SDO	Standards Development Organization
SEMI	Scientific/Engineering Motion Imagery
SMPTE	Society of Motion Picture and Television Engineers
SPI-3D	Standoff Precision Identification in Three-Dimensions
SRT	Secure Reliable Transport
ST	Standard
TCP/IP	Transmission Control Protocol/Internet Protocol
TRE	Tagged Record Extension
TS	MPEG-2 Transport Stream
UDP	User Datagram Protocol
UHD	Ultra-High Definition
VANC	Vertical Ancillary Space
VMTI	Video Moving Target Indicator
VNIIRS	Video National Imagery Interpretation Rating Scale

Appendix E MISB Documentation

E.1 Standards (ST)

A document is eligible to be a Standard when it meets at least one of the following criteria:

- Facilitates interoperability and consistency
- Defines metadata items

A MISP standard (ST) mandates binding technical implementation policy. As such, Government procurement actions should reference them as a mandatory conformance item for vendor offerings for Government acceptance.

For point of clarification, in commercial practice most identified standards (notably those from SMPTE) are "voluntary" standards, where equipment manufacturers and users are free to choose to comply or to not comply with the standard. Standards, as represented in this MISP, are not voluntary; they are mandatory.

A document begins the standards process as "Developing," and once authored subjected to community review and approval. Once adopted the developing standard moves to an "Approved" status. Standards that are obsolete or replaced are declared "Deprecated," while those no longer in use are "Retired" (see [109] for more information on the standards process).

E.2 Recommended Practice (RP)

An RP:

- Provides guidance that facilitates the implementation of a standard
- Typically, nonbinding but when used states requirements for its usage

Recommended Practices provide technical implementation policy. As such Government procurement actions may reference them as a mandatory conformance item for acceptance of vendor offerings.

E.3 Technical Reference Material (TRM)

A TRM is an informative/educational document that does not contain requirements. A TRM may result from a study or provide additional background to practices promoted by other guidance.

E.4 Engineering Guidelines (EG)

The MISB does not support the publication of new EG documents. An EG which meets the appropriate criteria listed above promotes to a ST or RP, while other legacy documents continue to carry this designation.

Appendix F MISB Processes

To facilitate the development of MISB ST and RP documents the following guides are available on the MISB website.

F.1 Document Development Process

The Document Development Process [109] defines procedures to be followed for: submission of candidate standards, the election of co-authors and technical assistance, the process of review/approval, and final publication.

F.2 Author Kit

The MISB Document Author Kit [110] provides a Microsoft Word template for developing a MISB document.