WHITE PAPER ON DISSEMINATING MOTION IMAGERY OVER CONSTRAINED NETWORKS

FRANCOIS URBAIN
JAMES MCAVOY, CD, BSC, MCS

BACKGROUND

In past decade, the Canadian Armed Forces have deployed video sensors on various stationary (i.e. towers) or mobile platforms (i.e. UAV, aircrafts, ships, armoured vehicles, etc.). All these sensors improved our troops and commanders' *intelligence, surveillance and reconnaissance* (ISR) capability. To ensure video can be viewed, shared and processed, sensor systems must adhere to NATO standards that enforce the transport and storage format of digital *Motion Imagery* (MI). In context of defence systems, MI is video combined with metadata, where coded information such as capture time, sensor and target location, platform positions and security classification are related to frames (i.e. images) in the video sequence as illustrated in Figure 1. This metadata with video enhances our ability to process, exploit and disseminate a vital intelligence product within national and coalition ISR systems.

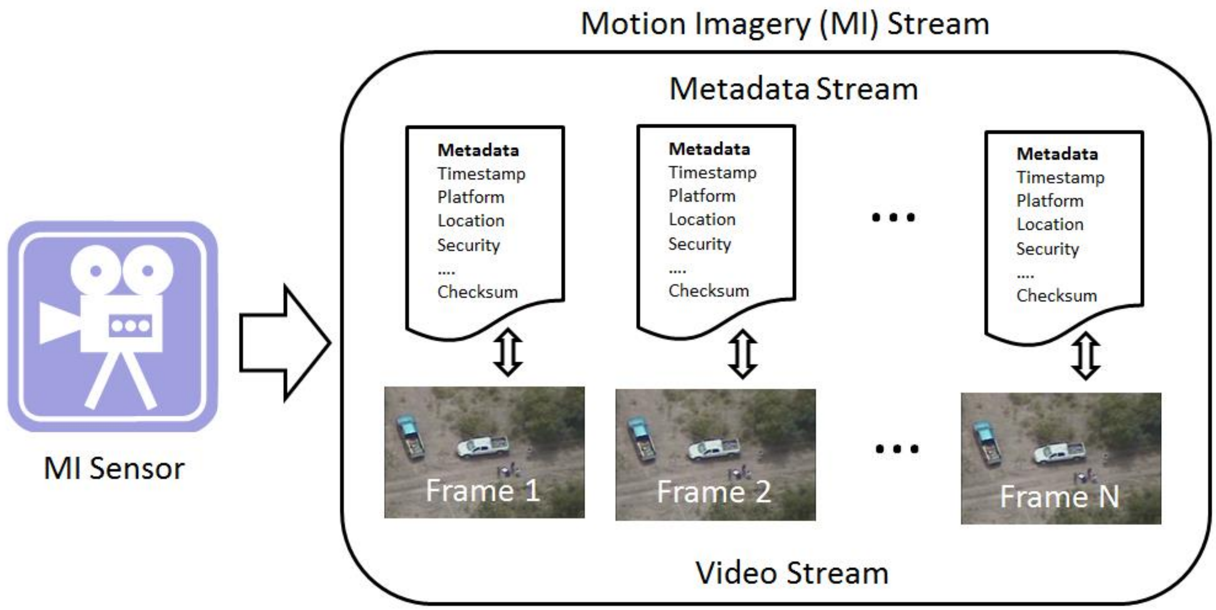

Figure 1. Motion Imagery (MI) is video stream infused with metadata, where metadata is associated with a frame in the video sequence.

THE PROBLEM

Streaming Motion Imagery over wired or wireless networks consumes tremendous amount of bandwidth. Sometimes the raw MI product transmitted from the sensor will exceed what a data network can bear. In the past several years, *Director Land Command Systems Program Management* (DLCSPM) has been developing solutions to distribute real-time MI over constrained networks. These solutions adhere to standards to ensure the greatest level of interoperability with national and coalition ISR systems. Following STANAG 4609, "NATO Digital Motion Imagery Standard", is the key enabler to guarantee this interoperability. The goal of this white paper is to provide a high level overview of a solution to disseminate MI over constrained networks, which is a fielded and sustained capability within the Canadian Armed Forces since the late 2000.

THE SOLUTION

In the mid-2000, DLCSPM recognized the issue of streaming bandwidth intensive Motion Imagery over constrained networks when deploying MI systems to *Forward Operating Bases* (FOB) while in Afghanistan. It was observed that wired networks provided sufficient bandwidth to distribute real-time MI, but wireless network fell short. There are several strategies one can employ so video can be transmitted over low-bandwidth networks, such as video down-sampling which decreases the number of bits require to code the video. Software and hardware products exist in the commercial and open source domain to do video down-sampling. However, none of these products reduce and/or process the embedded metadata stream in the MI product. Some applications would stream down-sampled video without any metadata. Others would transmit a MI product where the metadata stream consumed more bandwidth than the down-sampled video sequence.

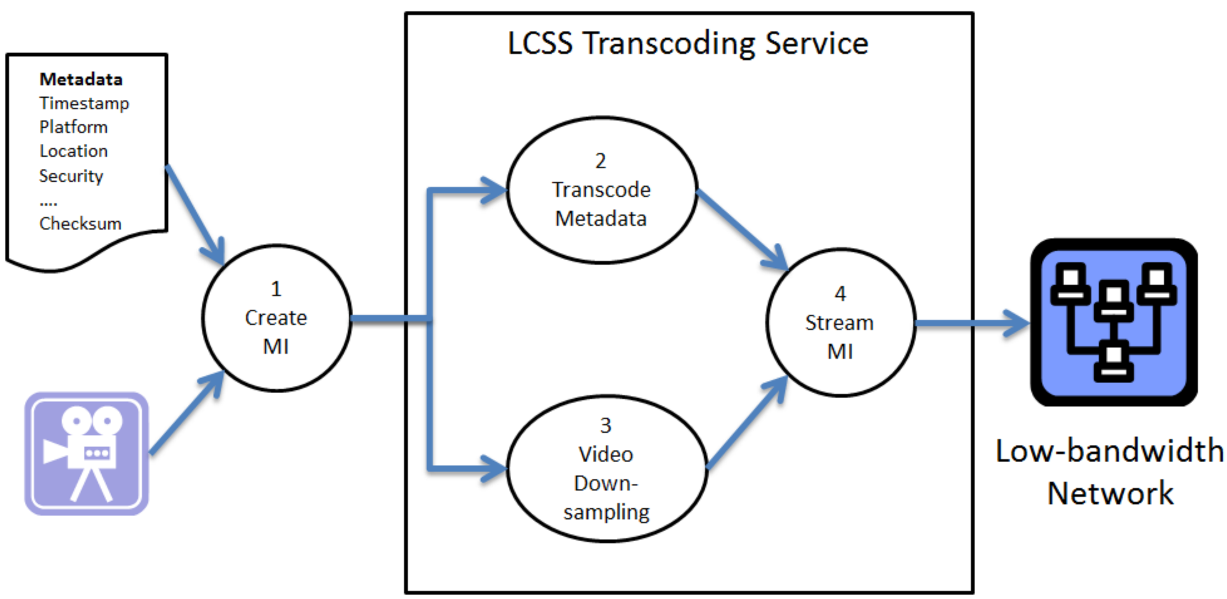

Figure 2. A block diagram displaying the steps to generated a down-sampled STANAG 4609 compliant MI stream.

Since no existing product on the market could deliver what was needed at the time, DLCSPM decided to develop a solution. In the following subsection, we will cover the steps to down-sample a STANAG 4609 compliant MI product so it can be transmitted over constrained networks.

STEP 1. CREATE MI PRODUCT

In most case, the sensor system transmits a compliant STANAG 4609 product, which is the NATO standard governing Digital Motion Imagery. Systems adhering to this standard will have a higher degree of interoperability with *Land Command Support Systems* (LCSS) applications and with other NATO ISR systems. There are occasions where sensor systems produce a non-conforming MI stream. In these cases, DLCSPM have developed applications to create a compliant STANAG 4609 stream. In general, this step maintains the video original signal (i.e. no down-sampling) but infuses metadata with the video to create a compliant STANAG 4609 MI product.

STEP 2. TRANSCODE AND DOWN-SAMPLE METADATA

Metadata can be binary encoded in various manners. Even within STANAG 4609 provides guidance on different metadata encoding methods. Within LCSS, it was decided to select metadata encoding standard that is based on MISB ST 0601, “UAS Datalink Local Metadata Set”. It was chosen because its popularity in ISR and UAV community for being a robust, simple and bandwidth friendly method to encode metadata. Over the past decade, numerous successful NATO and Canadian MI systems have implemented this standard to facilitate interoperability.

A unique capability with our Canadian system is the ability to modify and down-sample a metadata stream contained in a MI product in real-time. This feature allows us to edit or extend the metadata content, such as changing or adding the security classification to a MI product. Down-sampling ensures that the metadata stream is resized to save bandwidth without degrading the information.

STEP 3. VIDEO DOWN-SAMPLING

Down-sampling is the process of reducing a video’s bitrate so it can meet imposed storage or network constraints. Systems employ several different strategies and techniques to decease bitrate, such as downsizing frame resolution, lowering frame rate, and/or increasing quantization level. All these techniques and many more an application can combine to deliver the desired bitrate. LCSS leverages existing *Open Source Software* (OSS) products to provide this functionality. However, video down-sampling does decrease image quality, which impacts interpretability. Therefore, image quality is sacrificed as trade-off to bitrate. Determining the optimal balance between bitrate and quality is a challenge.

STEP 4. STREAM MI OVER CONSTRAINED NETWORK

Step 2 and 3 generate a media stream: one containing metadata and one containing video. These two streams need to be combined to create a STANAG 4609 compliant MI product that can be transmitted over a low bandwidth networks. The application that does this process is called a *multiplexor*. DLCSPM developed this application and has been leveraged in other systems that required similar capability to combine metadata and video to generate a STANAG 4609 compliant MI stream for transport and/or storage.

SOLUTION IMPLEMENTATION

Except for Step 1, the solution manifest in a DLCSPM application called the *LCSS Transcoding Service*. Figure 2 shows which steps are implemented in this application. By default, the service defines two down-sampling profiles: **Medium** and **Low**, which are outlined in Table 1.

Table 1. The table displays LCSS Transcoding Service default down-sampling profiles.

Down-Sampling Profile	Video Bitrate (Kbps)	Video Resolution Scaling	Video Frame Rate	Video Compression Standard and Profile	Metadata Frequency	
					Dynamic	Static
Medium	512	Half (1/2)	15 frames/sec	H.264/AVC Constrained Baseline	4 Hz	1 Hz
Low	64	Quarter (1/4)	15 frames/sec	H.264//AVC Constrained Baseline	1 Hz	15 Seconds

The Medium profile target bitrate is 512 Kilobits per second (Kbps). This profile cuts the original video resolution by half and reduces the number of frames per second (fps) to 15. The down-sampling profiles' frame-rate is approximately half what a North American television refresh rate of 29.97 Hz. The service encodes the down-sampled video using H.264/AVC, which is the most popular video compression standard on the Internet because of its greater compression efficiency over older standards. To ensure low coding complexity and latency with increase error resiliency, down-sampled video stream conforms to H.264/AVC Constrained Baseline profile.

For very narrow bandwidth networks, the LCSS Transcoding Service offers a Low down-sampling profile that has a target bitrate of 64 Kbps. The original video resolution is resized by a quarter. The other down-sampling attributes are the same as the Medium profile.

Table 1 shows metadata is separated into two categories: **Dynamic** and **Static**. Dynamic metadata is information that changes frequently such as platform and target location, timestamp and platform heading to name a few. Static metadata is information the usually remains constant during streaming such as sensor and platform name, security classification, release instructions, mission id, etc. By dividing the metadata into two categories and transmitting it less often, we were able to reduce the metadata bitrate by more than 50% for the medium profile and 90% for the low profile.

If the default down-sampling profiles fail to meet a client's requirements, the LCSS Transcoding Service provides a number of pre-defined profiles. Also, a technical support staff can extend the list of pre-defined down-sampling profiles via text file without rebuilding the application.

In the next section, the paper will show examples when the Medium and Low down-sampling profiles are applied to a video sequence.

APPLIED EXAMPLES

This section shows the effects when applying the default LCSS Transcoding Service Medium and Low down-sampling profiles on a *High Definition* (HD) video test sequence that has 10 seconds play length. Table 2 provides a summary of the original video attributes as well as the Medium and Low down-sampled video sequence. Table 2 also shows the calculated compression ratio when either Medium or Low down-sampling profiles are applied. In this example, the Medium profile was able to reduce the original video sequence by more than 90%, while the Low profile was able to reduce it an astonishing 99%!

Table 2. Table outlines the attributes for the original, Medium and Low profile video sequence.

Profile	Resolution@framerate	Bitrate	Size	Compression Ratio
Original	1920x1080@50	6 Mbps	7,789 KB	0%
Medium	960x540@15	512 Kbps	741 KB	90%
Low	480x270@15	64 Kbps	111 KB	99%

Of course nothing is for free, especially in video compression. The more compression is applied, the more distortion is introduced (i.e. poor image quality). The following subsection shows visually the effects of applying Medium and Low down-sampling profiles on one frame in a HD video sequence.

ORIGINAL VIDEO

The original is a very high quality HD video sequence, which Figure 3 shows one frame in that sequence. Figure 4 shows a zoomed in an area in that frame. You will notice that you can distinguish cars and people in the scene, which is more noticeable when video is playing.

Figure 3. The picture shows a frame of the original HD video sequence. The white outline box marks the zoomed in area.

Figure 4. The picture shows a zoomed in area in the original frame displayed in Figure 3.

MEDIUM PROFILE

When Medium profile is selected, the down-sampled video resolution is halved. Even though the medium profile compressed the original video by more than 90%, the overall subjective image quality is very good. However, if you zoomed in the same area displayed in Figure 4, you will notice a loss of some information. In Figure 6, the people that were obvious in Figure 4 are difficult to interpret or gone. Playing the video actually improves interpretability because objects moving in the scene can be inferred.

Figure 5. The picture shows a frame down-sampled using Medium profile.

Figure 6. The picture shows a zoomed in area of frame displayed Figure 5.

LOW PROFILE

Applying the Low profile, the original video sequence is compressed more than 99% but you will pay with greater image distortion (i.e. poor image quality). The video sequence transmitted using this profile still has value by providing basic *Situational Awareness* (SA) capability. If you were able to play the video sequence, you will still be able recognized large features in the scene, such as buildings, roads, boat and cars. The zoomed in area displayed in Figure 8, it is difficult or nearly impossible to recognize cars or people. Again, playing the video improves interpretability.

Figure 7. The picture shows a Low profile down-sampled frame.

Figure 8. The picture shows a zoomed in area of the frame displayed in Figure 7.

SUMMARY

Canada is deploying more Motion Imagery (MI) systems in domestic and overseas operations. Metadata embedded in a MI stream with video enhances ISR processing, exploitation and distribution. Sharing this vital ISR product over constrained networks is an issue. DLCSPM has developed solutions to facilitate the MI dissemination over low bandwidth networks. To ensure that the solution will interwork with existing and future ISR systems within Canada and with NATO partners, DLCSPM enforces the compliance with standards. STANAG 4609 is the NATO standard that mandates the format of digital Motion Imagery for transport and storage. The paper outlines a solution based on two guiding principles that influences MI systems development:

- Video with metadata enhances ISR functions; and
- Complying with standards ensures interoperability.

The DLCSPM application, LCSS Transcoding Service, encapsulates the solution and principles presented in this paper. The service provides two default down-sampling profiles: Medium with a target bitrate of 512 Kbps and Low with a target bitrate of 64 Kbps. Both profiles stream STANAG 4609 compliant MI product that embeds the metadata with the video. Since late 2000, the LCSS Transcoding Service has successfully delivered a capability to share MI products with national and NATO ISR systems over constrained networks.

Francois Urbain occupies DLCSPM 4-PMO ISTAR C2 position for the last five years and has delivered numerous ISR systems.

James McAvoy, C.D., B.Sc., M.C.S. has been providing engineering services to DLCSPM 4-PMO ISTAR C2 for the last five years by developing software to enhance tactical ISR capabilities. James graduated with Master of Computer Science from Carleton University studying in the area of video processing and compression.