Intelligence Automation Using WAMI for Counterinsurgency Applications

Ryan Martens
PV Labs
175 Longwood Road South, Suite 400A
Hamilton, ON Canada L8P 0A1
rmartens@pv-labs.com

ABSTRACT

The collection of persistent Wide Area Motion Imagery (WAMI) provides many key benefits in Counterinsurgency/Counterterrorism (COIN/CT) applications, achievable primarily through the contextual view with which fusion of other sensor data may be performed. Persistent WAMI provides situational awareness and facilitates sensor cross-cueing and sensor resource management. In addition, the forensic aspect of a persistent surveillance system allows new information to be obtained a-posteriori, or after an event of interest has occurred. Using this information, it is often possible to gain new insights on previous events, possibly using forensic data from other sensors, if available. It is also possible to analyse patterns of life over a period of time which leads to gathering of additional intelligence.

Collection of persistent WAMI, however, poses increasing challenges for analysts working with data acquired by such systems. As sensors migrate from tens to hundreds of megapixels to gigapixels in scale, exploitation of information from imagery is becoming increasingly difficult. The inability of an analyst to monitor an entire data set simultaneously drastically changes the concept of operations when using such a system. Additionally, persistent surveillance systems add a dimension of time, and also add the ability to alter the time axis at random depending on the occurrence of events of interest.

A WAMI solution therefore needs to incorporate exploitation tools and analyst efficiency as key requirements in its design, weighted against processing, storage, and dissemination bandwidth constraints. While spatial and temporal resolution have traditionally played a critical, and in many cases, sole role in system requirement definition, they fall short in many respects. Other factors, including image segmentation, real-time availability, sensor fusion, data access interfaces, and data abstraction models, are essential in simplifying the intelligence extraction from WAMI as well as WAMI combined with other persistent sensor technologies. A common spatial-temporal database for storing geo-referenced, time-tagged data of any format from a multitude of sensors is a key requirement for sensor fusion. The parallel processing architecture provided by modern GPGPU technology is instrumental in achieving real-time availability and a scalable data throughput. By shifting the focus to information extraction, analysts can be provided with the tools to extract intelligence for COIN/CT applications.

In order to fully automate the task of collecting intelligence from imagery, human analysts must be relieved of many of their tedious and monotonous responsibilities. An approach is presented which enables automated tracking from airborne platforms for determining geo-registered tracks for thousands of targets across the large areas of coverage provided by WAMI sensors. Traditional tracking algorithms often struggle when dealing with obstructions in visibility and lack of motion in the target of interest. An approach is discussed which leverages key properties of the images obtained, rather than tracking of the detections themselves, improving efficiency, robustness and reliability of the tracks obtained. In addition to improving overall throughput achievable by an analyst, this also enables an analyst to pursue higher-level intelligence gathering.
tasks which focus on fusion of information, rather than simply identifying and following a target of interest.

1 THE CHALLENGE OF MORE PIXELS

Recent technological advances have created a significant trend in Intelligence, Surveillance and Reconnaissance (ISR) applications towards persistent surveillance, that is, the collection and storage of information which can be converted into actionable intelligence in real-time or near real-time by maintaining enduring contact with the target or targets of interest [1]. Alternatively, persistent surveillance is a collection technique emphasizing the real-time ability of a system to detect, locate, identify, track, and target an object or objects of interest while lingering over an area. The ability to collect and store imagery persistently has enabled analysts to increase the quantity of intelligence gathered from imagery. Sensors which collect imagery over larger and larger geographic areas, as well as platforms containing multiple sensors of varying types both contribute to a significant increase in pixels that must be processed, stored, and analyzed to convert sensor data into actionable intelligence.

1.1 WAMI

Wide-Area Motion Imagery (WAMI) is produced by ISR assets capable of producing imagery from sensors approaching 100 megapixels to one or more gigapixels in scale. Frame rates are typically lower than traditional full-motion video (FMV) sensors, but usually operate at a minimum of one Hertz [2]. The resultant system provides high-resolution imagery over large geographic areas for long periods of time, allowing surveillance over city-scale regions, and enabling intelligence to be gathered on motion patterns and locations of many simultaneous targets over a large region of interest. The unique abilities of WAMI sensors provide contextual information unattainable using narrow field-of-view (NFOV) or “soda-straw” sensors. This information, collected over extended periods of time, provide analysts with the data to develop models of patterns-of-life, or gather Activity Based Intelligence (ABI).

WAMI is also commonly referred to as Wide-Area Persistent Surveillance (WAPS), Wide-Area Aerial Surveillance (WAAS), Persistent Wide-Area Surveillance (PWAS), and Large Volume Streaming Data (LVSD). Sensors capable of collecting WAMI data consist of an array of focal planes combined with algorithms for mosaicing, stabilizing, and georegistering [2] the collected information into a single, unified “virtual” focal plane. These sensors are mounted on airborne platforms capable of orbiting over a fixed location, and must be capable of pointing the sensor at this fixed location for extended periods.

The benefits of WAMI, however, do not come without cost. The large size of each image frame imposes significant constraints on system bandwidth for transmission and processing of the data, and enormous constraints on the storage subsystem, which must deal with terabytes of data, both as the data is collected, and as analysts and client applications extract intelligence from the collected imagery. Table 1 summarizes bandwidths and storage requirements for various motion imagery resolutions and frame rates. Note that these are raw (uncompressed) values. The Size, Weight, and Power (SWaP) of such systems often result in cost-prohibitive designs. With constant gains in processing capabilities, however, driven by developments in the consumer electronics industry, system SWaP is decreasing, increasing the affordability of WAMI technology.

The current workflow encompasses the Processing, Exploitation, and Dissemination (PED) of raw imagery into intelligence useful to the warfighter [3]. The large data volumes and processing required for working with data of this scale create challenges for all aspects of the PED chain. It is clearly not feasible for a single analyst to monitor the data collected from a single WAMI system. At a resolution suitable for intelligence
gathering, there is simply too much information for a single user to process and monitor. While multiple analysts could ultimately each monitor a subsection of the imagery, other improvements in workflow or efficiency are clearly required. For comparison, a single combat air patrol (CAP) for a Reaper or Predator providing 24/7 coverage over a given region of interest consumes up to 30 analysts for exploitation of standard FMV motion imagery [3]. The size of team required to support a platform containing a WAMI sensor as well would grow considerably.

Automation of many of the more mundane tasks could free analysts to perform higher-level analysis and intelligence gathering. Tracking, in particular, has the potential to fully automate or semi-automate many cognitively demanding tasks. When applying tracking to WAMI, however, a variety of challenges exist. In addition to the large data volume and processing requirements, mentioned previously, tracking applied to WAMI is also negatively impacted by the low frame-rate and reduced spatial resolution of the acquired imagery [4].

<table>
<thead>
<tr>
<th>Type</th>
<th>Resolution</th>
<th>Bandwidth</th>
<th>Storage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Television</td>
<td>640x480 @ 30fps</td>
<td>9 Mpixels/sec</td>
<td>39 GB/hr</td>
</tr>
<tr>
<td>High-Definition TV</td>
<td>1280x720 @ 24 fps</td>
<td>21 Mpixels/sec</td>
<td>74 GB/hr</td>
</tr>
<tr>
<td>Medium-Res WAMI</td>
<td>120 Mpixel @ 2 fps</td>
<td>240 Mpixels/sec</td>
<td>844 GB/hr</td>
</tr>
<tr>
<td>High-Res WAMI</td>
<td>1.5 Gpixel @ 5 fps</td>
<td>7.5 Gpixels/sec</td>
<td>26.4 TB/hr</td>
</tr>
</tbody>
</table>

Table 1: Bandwidth/storage requirements for various video sources. All signals are assumed uncompressed monochrome and 1 Byte/pixel.

1.2 Data Silos

With enhanced focus on ISR solutions, the explosion in the quantity and the variety of sensors has created an integration challenge. While standards are becoming more and more prevalent, the existence of legacy systems and inadequacies of integration tools has pushed the responsibility of data fusion to the analyst. A side-effect of this is that true fusion of data from various sources occurs significantly down-stream in the processing pipeline, typically, after the airborne platform has landed. Real-time integration of data from various sources, to this point, has been limited in scope.

In addition, the extra hardware required for each sensor’s storage and processing significantly contributes to the SWaP of fielded multi-INT solutions. Clearly, if decisions are to be made based on fusing data from multiple sources, tighter integration between various solutions is required.

The situational awareness provided by WAMI contributes to another benefit on platforms with multiple sensors. Cross-cueing of sensors, or tipping and cueing, is the action of targeting a slave sensor based on the output of a master sensor. A challenge with NFOV sensors is locating a target through the soda-straw view of the world. The contextual view provided by a WAMI sensor enables an analyst, or for that matter an automated algorithm, to locate a target of interest and cue the NFOV sensor to the desired location. Clearly, a long-term goal is a platform with multiple sensors autonomously detecting targets of interest, and sharing this information to cue other sensors, maximizing the benefit of the system. Sensor management, cross-cueing, and data fusion enable the most optimal sensor to be located in the right place at the right time.
2 COIN/CT

The United States’ 2010 Quadrennial Defense Review placed significant emphasis on Counterinsurgency (COIN), Stability, and Counterterrorism (CT) operations, with an expansion of ISR assets and improved intelligence, analysis, and targeting capacity [5]. This includes rapid processing, exploitation and fusion of collected data. A specific challenge of COIN/CT operations is the wide range of scenarios experienced in theatre, as well as the coordination of multiple resources and techniques in arriving at a solution. Densely populated urban areas with high concentrations of civilians increase the complexity of the operations and intelligence gathering. COIN/CT applications require a major shift from traditional ISR approaches. Instead of targeting large, easy to spot formations, intelligence must be gathered on individuals, who are typically much more difficult to locate [6].

The primary benefit of WAMI for these use cases is the added situational awareness and contextual information provided by the wide field of view and continuous coverage over time. Since COIN operations are network and relationship-focused, WAMI provides a key tool in gaining a true understanding of the insurgent’s environment, such as how groups are organized. Intelligence is a key component in COIN, as opposed to kinetic force. COIN requires a thorough understanding of the local population which defines the operational environment. Furthermore, intelligence must be shared across a large number of participants, not all of them military. The contextual view provided by WAMI systems provides a framework against which other intelligence data may be referenced. Intelligence sources, on their own, can provide misleading information if taken out of context; WAMI provides this context and reference point, supporting fusion of Signals Intelligence (SIGINT), Human Intelligence (HUMINT), and other sources of intelligence.

A less-obvious role of ISR for COIN applications is its non-intrusive nature. Traditional operations have a disruptive effect, resulting in temporary changes during the course of the disruption, but returning to insurgent control once the disruption has ceased [7]. WAMI systems are able to collect intelligence over increasingly large areas, often without any knowledge of the sensor’s presence.

Three dominant modes of operation together provide significant operational advantage for COIN/CT: forensic, tactical/overwatch, and predictive [8]. Forensic application allows analysis or modelling to be performed following an event of interest by monitoring events and other activities surrounding the event. Overwatch provides real-time, tactical information; situational awareness and contextual information assist in this mode of operation. Finally, predictive applications support the development of models and analysis capable of detecting anomalous behaviour and preventing the occurrence of an incident.

3 PUTTING THE PIECES TOGETHER

Development of a WAMI system for COIN/CT applications requires careful design of the sensor, and a corresponding system architecture which enables actionable intelligence to be extracted from the large data sets quickly and efficiently. This discussion will not focus on the traditional design parameters of spatial and temporal resolution. While a requirement for high-resolution is implicit in the definition of a WAMI sensor, other factors contribute to the utility of the information acquired as well as the ability to apply the PED workflow to extract meaningful intelligence, either by an analyst, or autonomously. In addition, a requirement for high temporal frequency, or frame rate, has been driven by the adoption of FMV sensors, which operate at very high frame rates. Traditional tracking approaches work very well with high resolution, high frame rate data. By revisiting the approach employed for tracking, and carefully examining other aspects of the system design, discussed in the following sections, a path is presented for a manageable and realizable WAMI
3.1 Image Segmentation

A basic requirement of a WAMI system for intelligence automation is image segmentation. The scale of a single frame is orders of magnitude greater than frames in traditional FMV applications. Information exploitation performed on this scale of data increases processing, networking and storage requirements substantially, as well as the time to perform required operations. For COIN/CT applications, there is a clear benefit in being able to work with subsets of imagery, pertaining to specific Regions of Interest (ROI).

In a simple implementation, a chip-out allows a client application or analyst to select a region of fixed or variable size, encompassing a subset of the high-resolution WAMI frame. The reduced processing burden enables real-time processing or user access to the ROI. An improved scheme supports the ability to vary the resolution of the ROI independent of the acquired data. For example, it may be of benefit to examine a larger ROI at a correspondingly lower resolution for contextual decisions to be made not requiring full-resolution data. An arbitrarily flexible scheme supports image segmentation at varying scales, from a thumbnail view of the entire sensor field-of-view (FOV) to full-resolution views of any ROI in the frame, or any view in between.

Wavelet compression algorithms, such as JPEG2000, have provided many benefits over standard JPEG or similar compression approaches. A significant limitation of JPEG2000, however, results when generating lower-resolution samples encompassing many tiled images covering the entire frame. If the storage system relies on separate files for each of these tiles, the access overhead associated with generating this ROI increases substantially. More intelligent schemes still suffer from the access times associated with retrieving data across large portions of disks storing archived information. JPEG2000 Interactive Protocol (JPIP) overcomes many of these deficiencies, providing a client-server approach for loading portions of a large JPEG2000 frame at varying resolutions [9]. Generation of these large frames works well for post-processed applications, but falls short in distributed, scalable, real-time applications.

Traditional image pyramid approaches [10] offer greatly improved performance as tiles of varying resolutions support greatly improved efficiency without sacrificing the flexibility of the analyst or client application. In addition, the pyramid approach easily supports distributed caching and storage of data in a seamless manner [11]. A draw-back of the image pyramid results from the increased storage costs of storing discrete resolution levels in addition to the high-resolution imagery. Common regular pyramid schemes result in a 33% increase in storage requirements. This relative increase in storage cost is easily offset through the introduction of even a modest compression ratio.

3.2 Real-Time Availability

A fundamental requirement for supporting tactical and predictive modes of operation is real-time availability of data. Large volumes of imagery must be processed and stored in real-time within the bandwidth constraints of the sensors, as well as the processing, networking, and storage hardware. Furthermore, the data must be accessible for exploitation and dissemination during collection in near-real-time to real-time, potentially by remote analysts, operators, or warfighters over data links such as TCDL.

A rudimentary approach to reducing the resource requirements for capturing WAMI content is to split the data acquisition into several parallel streams, typically on a per-focal-plane-basis, such that sections of the spatial extent are archived in completely separate storage units. Unfortunately, exploiting data stored in this manner requires the same amount of aggregate bandwidth exhibited by the acquisition system itself, essentially requiring a complex backend system with a significant amount of post-processing before it can be fully
utilized, which is most-likely non-real-time. Such legacy designs trade relative simplicity on the front-end for severe limitations on content exploitation.

A more feasible approach includes a mosaicing step which unifies the data from individual focal planes, processing using one of the image segmentation methods described previously. This step enables bandwidth-reduced access to any subset of the collected data, either spatially or temporally, in constant time for a given output resolution. Effectively, the data is indexed and stored in a georeferenced spatial-temporal database. Once the data is pre-processed appropriately, the bandwidth required for data exploitation is substantially reduced. It is true that a greater burden is placed on the acquisition system to perform this operation, but once complete, the data can be stored in a format which reduces all subsequent processing operations, while enabling real-time access to the acquired data. This greatly reduces processing requirements to access the data, and enables future advances in Real-Time Analytical Processing (RTAP) and supports moving data processing and intelligence extraction closer to the sensor acquisition source.

Advances in GPU and distributed computing technologies permit low-cost, highly parallelized solutions that are massively scalable. With hundreds of processing cores on a single die, GPUs are ideally suited for the numerical floating point calculations required for the high pixel density of WAMI data. Furthermore, developments in the CUDA and OpenCL frameworks have enhanced the flexibility of GPU-based development efforts. While not all algorithms can be implemented with this degree of parallelism, calibration and distortion corrections, orthorectification and georegistration, multiresolution image calculation, and certain forms of image compression are ideally suited for GPU-based architectures. Furthermore, the gaming market has driven the performance of GPUs to advance at an impressive rate, with benefits immediately transferrable to other applications. These benefits are realized as reduced SWaP and real-time performance, while maintaining a highly scalable architecture to handle WAMI data of increasing magnitude [12]. It is important to note, however, that the compression employed is a lossless scheme. This prevents the occurrence of image artifacts which may negatively impact the tracking performance.

A final critical component required for real-time processing and availability of WAMI data is a distributed system architecture. As a single, high-resolution focal plane generates large volumes of data, small computation clusters or distributed architectures are required. Distributed caching strategies may be employed to allow distributed storage and parallel access of data by one or more analysts or client applications. Despite this design, however, it is imperative that the stored data is accessible in a self-contained storage media. This requirement facilitates ground-based retrieval of the data for long term archival, analysis and dissemination.

### 3.3 Data Fusion

As discussed previously, data silos inhibit the ability to make decisions from the information acquired. The seamless fusion of data from various sensors overcomes limitations of individual sensors, while enabling additional information to be used to piece together more pieces of the puzzle. For example, visualization of a scene using different modalities, different viewpoints, or different focal lengths can provide significantly different information, even using sensors mounted on the same platform. In addition, the ability to correlate data from different sensors accurately in both space and time is critical for leveraging the power of multi-INT.

System architecture imposes significant constraints on the utility of a multi-INT platform. Consider the system in Figure 1. If data from all sensors are stored separately and the previously discussed architecture is not implemented, significant limitations are observed when analysis of data is performed on the ground over a data link. A copy of the data from each sensor must be streamed over the data link, where an ROI from each sensor may be fused to create the desired view for the analyst. Alternatively, the system in Figure 2 stores the acquired data from each sensor in a common geospatial-temporal database. This allows common operations to be performed on each sensor stream, selecting an ROI and only transmitting a subset of the data over the data link for analysis. Performing the fusion and processing on-board the aircraft will further reduce the bandwidth.
requirement.

In addition, improved efficiency is achieved using a common container structure for all sensor sources, as opposed to separate files, or even worse, separate storage units. This container construct reduces the operating system overhead of working with a multitude of different files.

Figure 1: Example Multi-INT System Architecture

Figure 2: Improved Multi-INT System Architecture
3.4 Data Abstraction and Interface Design

Data abstraction is a useful tool for hiding the sensor-specific implementation details from an end-user, analyst or exploitation tool. Processing of sensor data is greatly simplified using an abstract data methodology. Inclusion of an extra configuration step greatly complicates the processing, exploitation and dissemination of the acquired data, resulting in more errors, greater access time, and increased complexity in the tools accessing the data. An abstract data model supports re-use of common toolsets across all data sources, simplifying workflow, and reducing time and cost of integrating new sources.

In addition to an abstract data model, a standardized access interface is required. The developed interface should not require an end-user to understand or be aware of the physical sensor configuration in order to work with the data and should support the abstracted data model. Furthermore, incorporation of standards, where appropriate, can greatly aid in achieving interoperability with other systems. It should be noted, however, that where use of a standard complicates the design for internal interfaces, an interface layer can be used to adapt external interfaces to the appropriate standards. In all cases, an interface layer increases flexibility by allowing any legacy component to interface with a system, even if that system does not conform to existing standards.

A simple model and interface abstracts the data access methodology to a camera placement in a common reference frame. For example, by positioning and orienting a virtual camera in an Earth-fixed reference frame, and including an appropriate FOV or bounding box, it is a simple operation to populate a buffer with georegistered image data at the required resolution, as shown in Figure 3 and Figure 4.

![Figure 3: Camera Location and FOV Specification](image)
3.5 Platform Considerations

The sensor platform is a major contributor to the overall performance of a WAMI system. Not only does the platform stabilization contribute to the overall quality of the image, but the geo-pointing accuracy is a major contributing factor in the system’s ability to georectify the acquired imagery. Furthermore, the number and parameters of the steerable axes of the pointing system determine the persistent area of a given sensor footprint. A well-designed platform solution can leverage the high-performance Inertial Measurement Unit (IMU) required for stabilization to accurately measure the sensor’s Line of Sight (LOS) in inertial space, which is used in the georegistration process [13]. If the Inertial Navigation System (INS) is synchronized between multiple sensors on the airborne platform, the relative accuracy of the georegistered solutions for each sensor will be significantly improved, enhancing the ability to fuse data accurately and efficiently.

Stability can potentially refer to three separate components: LOS jitter stability, GEO steering stability, and INS error stability [13]. LOS jitter contributes to pixel smear resulting from sensor movement during the integration time. GEO steering stability may cause the LOS to wander around the desired target location, reducing the persistent area. Finally, the INS error will cause small drift in the orthorectification of the acquired imagery.

As an example, for a system with a pixel pitch of 9µm and a lens with a focal length of 135mm, a 10ms integration time requires a stability jitter of 16.65µR in order to keep pixel motion below half a pixel, preventing smearing [13]. Increasing the integration time, which is desirable in order to increase the dynamic range, requires an even lower stability jitter target. The PV Labs LDG, or Look-Down Gimbal, is capable of achieving a stability jitter below 5µR RMS, significantly improving the potential to acquire imagery of a high quality for analysis.
3.6 Sensor Considerations

Another major consideration in a WAMI system is the sensor design. As discussed previously, a typical configuration for a WAMI payload design consists of an array of focal planes positioned to create a single unified virtual focal plane. This task is accomplished using algorithms for mosaicing, stabilizing, and georegistering. While this task can, and often is, performed in software, a process for calibrating individual focal plane positions and orientations greatly simplifies this task, significantly reducing processing requirements. A highly calibrated optical/sensor assembly allows other corrections to be accounted for and performed with great ease. PV Labs’ PSI Vision sensor payload has eight and twelve focal plane configurations, with resultant resolutions from 88 megapixels to over 300 megapixels. Figure 5 shows an eight focal plane configuration, with image seams highlighted for clarity.

Synchronized focal plane triggering and metadata collection are also critical for proper georegistration of the acquired imagery. A tight coupling between the platform’s INS and the payload allow for easy integration, by supporting a high degree of accuracy with camera triggering and instantaneous pose.

Also, gain configuration across all focal planes is imperative for automated algorithms. Tracking, in particular, is greatly affected by changes in focus and intensity at focal plane boundaries. Sensor Auto-Gain Control (AGC) algorithms running independently will cause a “checkerboard” or “quilt-like” effect in the acquired imagery, as the area covered by a WAMI sensor could have significant variation in illumination. In addition, the time-varying property of most AGC algorithms can result in sudden variations in intensity, generating either false or lost tracks.
4 AUTONOMOUS OPERATION

WAMI data was collected using a system designed according to the previously discussed guidelines, using PV Labs’ PSI Vision sensor and Tactical Content Management System (tCMS). Track generation applied to WAMI data, discussed in this section, was performed using the Tracking Analytics Software Suite (TASS) by Signal Innovations Group.

4.1 Tracking Applications for WAMI

As discussed previously, WAMI poses significant challenges for autonomous intelligence collection, in particular due to the large data volume and processing requirements, but also due to the low frame-rate and reduced spatial resolution of the acquired imagery. Some additional challenges, particular to tracking, are: 1) point-like moving objects, 2) parallax resulting from the orbiting platform generates motion clutter, and 3) inefficiencies in registration resulting from real-time approximations, whereby stationary objects may appear to move many pixels in distance [14].

Considering the wide array of challenges affecting the ability to autonomously extract intelligence from a WAMI system, it is instructive to examine the solution space as a spectrum. Furthermore, reluctance to accept the validity of an autonomous approach may impede adoption. Hence care must be taken, and solutions which simply offer tools or aids to a human user or analyst are of significant importance for preliminary users of a WAMI system.

A User-Driven Model has been proposed, whereby a computational system assists, rather than replaces, human users; this is accomplished by increasing the user’s effective situational awareness, and by automating the more tedious and labour-intensive tasks [15]. For example, a tool may operate in a fully manual mode, where an analyst augments acquired data with manual annotations or by manually selecting an object being tracked from frame to frame [11]. While tedious, this option eliminates errors that may arise from an analyst examining raw imagery unassisted, and provides an effective means of capturing the extracted intelligence for sharing with other analysts. Aids may be added to the tool to speed the process, but the basic operation is manual in nature. Extending this further, a tool may quickly provide a user with answers to simple queries, such as vehicles that have come or gone from a selected building, or all vehicles travelling above a certain speed. These semi-automated approaches greatly enhance an analyst’s productivity, while still leaving the analyst as the prime intelligence extractor and decision-maker. Analyst-aided tracking extends the level of automation even further, by allowing an analyst to guide the automated tracking algorithm when it encounters regions of difficulty. This approach has demonstrated the accuracy of a manual approach with the efficiency of an automated approach [16].

At the far end of the spectrum, a fully-automated approach allows intelligence to be gathered in an autonomous fashion. This approach enables activity-based intelligence (ABI), which is the study of moving entities, activities, patterns-of-life and networks [17]. Successful implementation of ABI relies on data-drive models; these models must ultimately be flexible and robust, as the patterns detected in large data sets are context-dependent and complex.

In the case of COIN/CT, the manual and semi-automated approaches are ideally suited for forensic and overwatch operations. Fully-automated approaches can be applied to any of the three operations, but are required for predictive use-case scenarios. Data-driven models applied to a fully-automated approach provide significant benefits in each of the three domains. In forensic applications, network analysis can determine links between players involved, both directly and indirectly, in an event of interest. In overwatch applications, the warfighter can be provided with real-time information including current locations of tracks of interest.
Finally, in predictive applications, anomalous behaviour, or behaviours indicative of a threat, can be detected prior to an event occurring, and used for deciding on a course of action [8].

4.2 Tracking Implementation for WAMI

The Tracking Analytics Software Suite (TASS), developed by Signal Innovations Group is a robust solution able to track many simultaneous objects over large regions of interest [16]. The solution includes an image registration and stabilization component, a dynamic object tracking component, various analyst and user interface solutions, and is capable of supporting multi-camera handoff. The scalable architecture supports processing of WAMI data of varying scale, with varying numbers of targets monitored. The dynamic object tracking component is based on a Bayesian framework. Employing the data-driven model approach, probabilistic models for background colour, foreground colour, object shape, and object motion are learned and adapted dynamically as new data is acquired [16]. Any uncertainty associated with each model propagates forward, allowing the exploitation of the true confidence in each generated track [16]. Finally, it is important to note that after a track has been detected, following frames do not perform a detection step with a corresponding data association. This behaviour supports robust operation, even in dense urban environments [17].

Figure 6: Track Display of ROI
It is important to note that while the described approach is able to operate successfully without sensor metadata or camera models, and is capable of processing separate focal planes individually, this complicates the problem and increases processing requirements. Since this in turn increases the overall system SWaP, this is undesirable. The aforementioned approach, with its stabilization, highly accurate metadata, and abstracted data access model creating a unified virtual focal plane, remove this requirement, resulting in a highly efficient solution.

The results presented are from a flight test over an urban area in Whitby, Ontario, Canada, conducted in May 2011 at an altitude of approximately 2,200 meters. Figure 6 shows the detected tracks over a small ROI. Note that tracks were calculated over the entire FOV; the ROI is shown for visibility of the tracked objects. For reference, Figure 7 displays all tracks covering the entire FOV, an area over two kilometers by two kilometers in size. In this region of interest, approximately 2,800 individual and unique tracks were detected and monitored throughout the flight.

Figure 7: Track Display Covering Complete Sensor FOV
4.3 Data Fusion and Multi-Sensor Platforms

As the previous sections demonstrate, RTAP is quickly becoming a reality. The scalable system design supports real-time tracking of WAMI data. In multi-sensor airborne platforms, data fusion enhances the utility of the fielded solution. Tracking performance can be enhanced making use of multi-modal solutions, and sensor cross-cueing can be employed to augment the intelligence extracted from the data acquired. As autonomous system operation becomes more commonplace, areas of study such as sensor resource management can be applied to improve the overall system efficiency. Sensor coverage can be improved using coordinated activity planning. Leveraging activity based intelligence on a multi-sensor platform can ensure that resources are tasking accordingly for maximal benefit. Models of patterns-of-life, track anomalies, and event detection anomalies can all provide intelligence of far greater valuable than an individual ISR asset is currently capable of acquiring on its own.

Furthermore, user interface requirements can be employed to aid an analyst in disseminating actionable intelligence. As Figure 8 shows, standards such as the KML file format, can be leveraged to quickly import track data into common tools, such as Google Earth. This clearly demonstrates the utility of extracting intelligence, rather than just extracting imagery.

![Figure 8: WAMI Track Display in Google Earth](image-url)
5 CONCLUSIONS

In conclusion, the previous approach defines a path for designing a WAMI solution as a system. Many of the design choices from the platform and stabilization, to the sensor payload, and processing architecture and methodology, bear significant impact on the ability to automate the conversion of imagery to actionable intelligence. In order to support the changing and often unknown demands of COIN/CT applications, a solution is required which can support forensic applications, overwatch or tactical operations, and also predictive analysis. An approach is presented supporting real-time generation of thousands of georegistered tracks as part of a complete WAMI solution. This approach leverages the platform, sensor payload and processing abilities of PV Labs’ LDG, PSI Vision, and tCMS components, combined with the TASS tracking solution from Signal Innovations Group. The approach successfully demonstrates the simultaneous tracking of thousands of objects across a large region of interest.

6 REFERENCES


[8] L. Kennedy, E. Wang; “Activity Recognition in Wide Area Motion Imagery”; National Meeting of the Military Sensing Symposium (MSS); July 2010.


