Video Surveillance and Monitoring Integrated Feasibility Demonstration Specification Document

Carnegie Mellon University
Sarnoff Corporation

May 18, 1997

Contents

1 Overview **2**

2 IU functionality **2**

2.1 Goals for Demo I IFD effort 2

2.1.1 MTD and tracking 2

2.1.2 Active sensor control and coordination 3

2.1.3 User interface 3

2.1.4 Activity monitoring 3

2.2 Goals for Demo II IFD effort 3

2.2.1 MTD and tracking 4

2.2.2 Active sensor control and coordination 4

2.2.3 User interface 4

2.2.4 Activity monitoring 4

2.3 Technology matrix 6

3 Demo I **6**

4 System architecture **7**

4.1 Operator control unit (OCU) 8

4.2 Sensor processing units (SPUs) 8

4.3 Synchronization of data and video 10

5 Multi-sensor coordination and calibration **11**

6 Data structures and communication **11**

6.1 Sensor data structures 12

6.2 Target data structures 12

6.3 Communication packets 12

7 Work Schedule **13**

1 Overview

This document presents the goals, capabilities, demo scenarios, and system architecture currently envisioned for the Video Surveillance and Monitoring (VSAM) Integrated Feasibility Demonstration (IFD).

The purpose of the IFD is to demonstrate a working prototype system for video surveillance and monitoring that:

1. provides real-time capabilities for cooperative surveillance by multiple air- and ground-based sensors
2. advances the state-of-the-art in IU functionality
3. favours video over still-image analysis

The IFD project is currently slated to run from May 1997 through December 1998, and will include two demonstrations. The first, Demo I, will take place 6 months into the project, in mid-November 1997. The second, Demo II, will take place in November 1998.

2 IU functionality

Four areas of functionality will be addressed:

- Moving target detection (MTD) and tracking
- Active control and coordination of sensors
- User interface that improves operator awareness of scene activity
- Activity monitoring

Specific target functionalities for Demos I and II are summarized below.

2.1 Goals for Demo I IFD effort

2.1.1 MTD and tracking

Target detection in the presence of

- camera motion
- varying target speed

- cessation of target motion

This will be accomplished through differencing stabilized imagery and through correlation matching.

2.1.2 Active sensor control and coordination

- Fixation of an operator designated area with a single airborne sensor to compensate for airplane movement and vibration.
  This will be accomplished through the use of frame-to-frame motion estimation coupled with active camera control.
- Multi-tasking of a camera between multiple designated areas so that an operator can monitor several regions simultaneously.
  The camera will automatically servo to aim at each area. Camera aiming that is robust to uncertainties in airplane position and attitude will be accomplished by using telemetry for coarse camera aim, and registering the video to a mosaic to refine the aim.
- Fixation of a moving target
  This will be accomplished by using detected motion to aim the camera.
- Ground-to-air target notification

2.1.3 User interface

The operator control unit (OCU) will present a graphical user interface that:

- Enables the operator to select regions of interest and perform calibration using an automatically-constructed aerial mosaic.
- Enables the operator to control algorithm parameters.
- Enables the operator to designate areas that should be monitored for incursion and areas in which motion should be ignored.

2.1.4 Activity monitoring

The demonstration will illustrate the use of the system for airborne detection of movement in selected areas and airborne verification of perimeter incursion detected from the ground.

2.2 Goals for Demo II IFD effort

Some of the goals listed below for the Demo II effort may be moved to Demo I if time permits.

2.2.1 MTD and tracking

- Robust detection and tracking of multiple targets in the presence of
  - motion of other objects / targets
  - varying target aspect
  - occlusion

  This will be accomplished by augmenting the motion- and correlation-based tracking algorithms with mechanisms for matching and disambiguation based on image or motion appearance.

2.2.2 Active sensor control and coordination

- Robust target tracking over extended ranges of movement using multiple ground and air cameras.
  This will be accomplished by matching targets between camera views on the basis of their position (coupled with calibration information relating camera fields-of-view) and appearance.
- Multi-tasking of a single airborne camera to track multiple widely-separated moving targets.
  This will be accomplished using the camera aiming algorithms developed for Demo I, coupled with prediction of target trajectories.

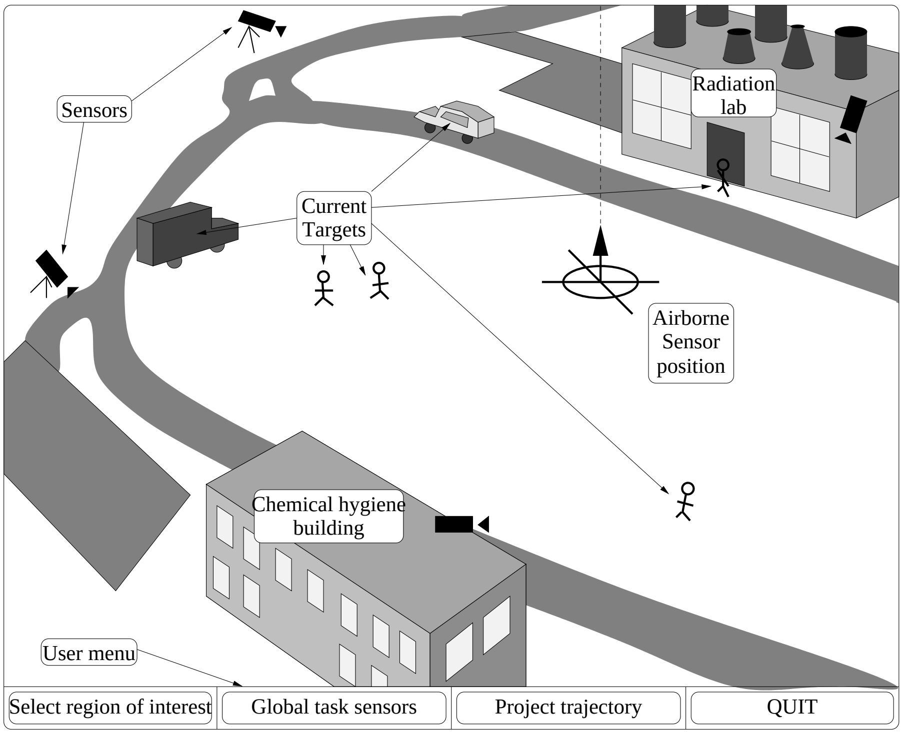
2.2.3 User interface

- Enables 3D visualization of site using DEM and CAD models.
- Enables operator awareness of each camera's position, field of view, and footprint.
- Enables operator awareness of targets and their previous locations.
- Enables operator to designate targets to be tracked or ignored.

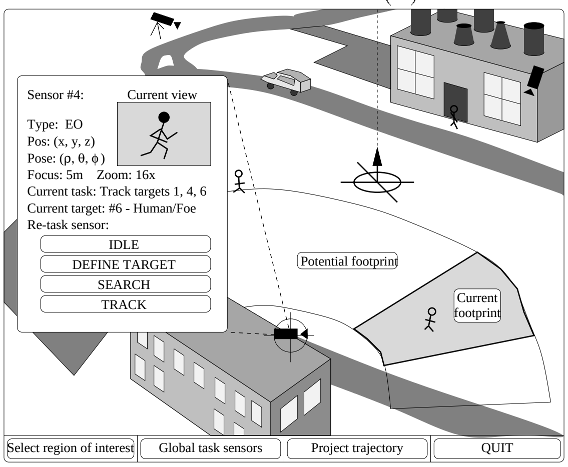
A potential user interface is shown in Figure 1.

2.2.4 Activity monitoring

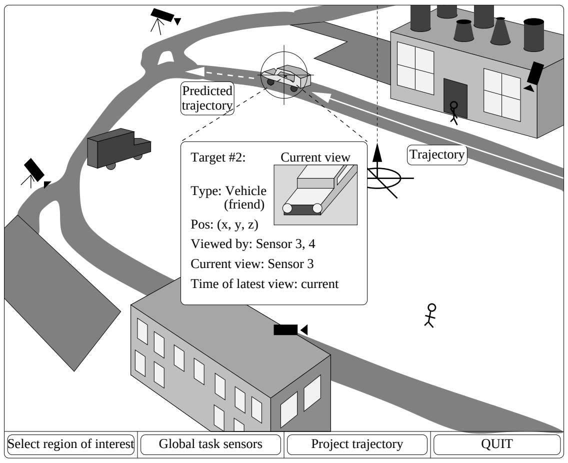
Identification of targets as rigid or non-rigid to allow discrimination between humans and vehicles.



(a) User interface components



(b) Available sensor data



(c) Available target data

Figure 1: Components of 3D user interface for the Bushy Run site

Capabilities	**Approach** *less sophisticated ⟷ more sophisticated*		
Scene Modeling	aerial mosaics	Digital Elevation Maps (DEMS)	VRML site models
Camera to Scene Registration	telemetry	pose, manual correspondence	pose, automated correspondence
Designating areas of interest	designate on mosaic	designate on site model	
Sensor Servoing	telemetry	visual landmark search	video to reference image registration
MTD & Tracking in a single view	temporal change	stabilization, temporal change	stabilization, temporal+correlation
Active Tracking	single target	multi-target	multi-target, fast moving
Target Location	within camera footprint	intersect view ray with DEM	multi-view triangulation
Visual Display	raw video with target indication	overlay trajectory on local mosaic	overlay trajectory on scene model

Table 1: Technology matrix for VSAM IFD

2.3 Technology matrix

Table 1 shows the technologies proposed by the IFD to accomplish VSAM tasks as the project continues. The minimum requirement for Demo I will be that it incorporates the technology described in the first column with subsequent technologies being developed and showcased in following demos.

3 Demo I

Demo I will demonstrate the above-described IU functionality within a military scenario. The scenario will illustrate the use of a VSAM system for airborne support of a covert operation at an enemy site, represented by Bushy Run. The scenerio will be as follows: An airborne platform (the Night Vision Lab's Islander aircraft) is flying above the site and monitoring multiple designated areas in which activity would be expected if the enemy became aware of the operation. Friendly commandos have dropped off a ground-based sensor in a hidden location to provide ground-based forward observation. The airborne sensor is cued by the ground-based observation of enemy movements of interest. The airborne sensor then tracks and monitors that activity.

A more detailed script for Demo I is as follows:

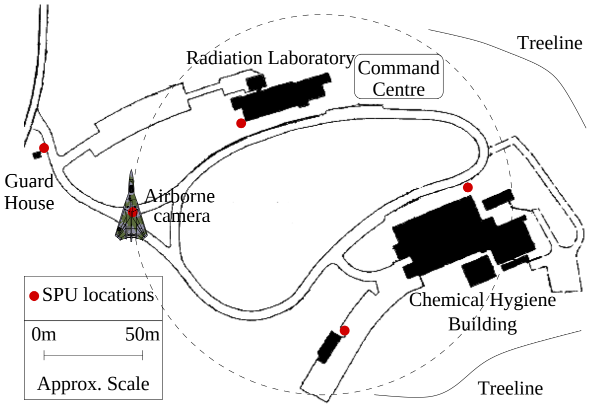


Figure 2: Setup of SPUs for year 1 demo

1. Islander flies over Bushy Run and generates an aerial mosaic (prior to the Demo). Mosaic is transmitted to the OCU and stored for use during Demo.

2. Human operator designates locations of interest on the mosaic (e.g. parking lot by Radiation Lab, and front door of chemical building).

3. Islander SPU pans to first point of interest and servoes, keeping it in field of view while plane moves. When operator clicks a button, camera is panned to second point of interest and servoes. This repeats. Potentially, more than two points of interest can be added, with multi-tasking switching from point to point in round robin fashion in response to operator request.

4. Suddenly, one or more forward observer SPUs detect a moving target that may be a threat to the operation. This information is transmitted to the OCU, which determines the geographic scene coordinates of the target. A command is sent to the Islander SPU to interrupt its multitasking and switch to the new target designated location. It acquires the moving target, and tracks it. If operator determines target is not a threat, then they can resume multi-tasking on the original points of interest.

4 System architecture

The VSAM prototype will consist of a single Operator Control Unit (OCU), coupled to multiple Sensor Processing Units (SPUs). Each SPU is connected to a sensor, e.g. a single camera, a stereo pair of cameras, or an omnidirectional camera. Video processing for detection and tracking for each individual sensor will take place at the SPU to reduce processing load at the OCU and to avoid problems with control loop latency and video transmission. Data communication between the OCU and the SPUs will take place over a (radio) Ethernet link using a communications protocol that will be published at a later date. Video will be transmitted from the SPU sensors back to the OCU over either coax cable or via microwave (or UHF) transmission.

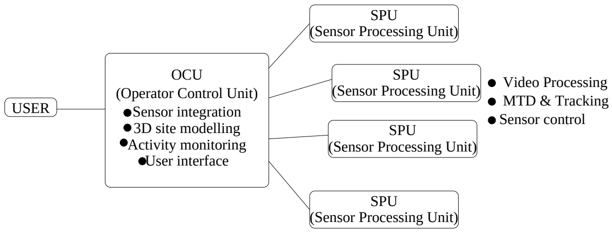


Figure 3: VSAM testbed overview

4.1 Operator control unit (OCU)

The OCU (Operator Control Unit) is the hub of the VSAM testbed. Its functions are to provide user visualisation, present a graphical user interface, represent targets in a scene coordinate system, and task SPUs. It will consist of two computers, one for graphics and one for coordination. The graphics engine will provide a 3D graphical user interface/visualisation tool. The coordination engine will task the SPUs and maintain the scene representation.

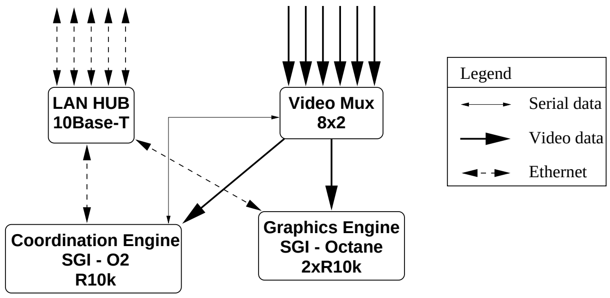


Figure 4: OCU configuration

Hardware components:

- SGI Octane 2xR10k/195MHz - Graphics display engine
- SGI O2 R10k/195MHz - SPU coordination engine
- Video Mux
- LAN Hub 12 port 10Base-T

4.2 Sensor processing units (SPUs)

SPUs (Sensor Processing Units) are sensor platforms capable of performing vision processing and IU functions such as moving target detection and tracking. The SPU consists of a processor responsible for video processing, sensor control, and communication with the OCU, together with a set of peripherals to allow such communication. The general components of an SPU are shown in Figure 5. Notes on each component follow below:

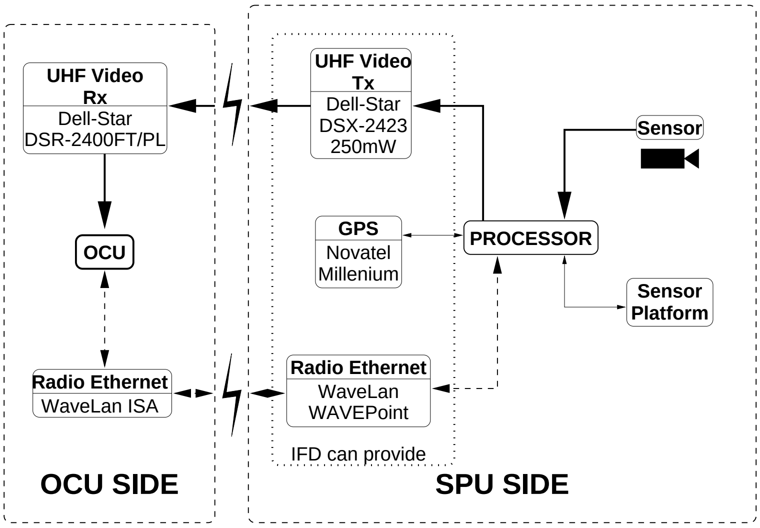


Figure 5: SPU Components

- Processor. Any processor desired by the IFD or FRE may be used, as long as it can communicate with the OCU via Ethernet by following the IFD OCU-SPU communications protocol. Typical processors envisioned include the IBM PC or Silicon Graphics O2, either of which may be augmented a Sarnoff VFE-100 image processor.

  Two additional requirements must be met. To enable the entire VSAM system to have a synchronized system clock that can be used to relate data and video obtained from different SPUs. First, the processor must be capable of outputting analog video that contains the last frame it has completed processing. This video will be transmitted to the OCU via the UHF link. Second the processor must allow a GPS device to be temporarily connected via a serial port so that the processor can be synchronized to the system clock by using the GPS timecode. (The IFD will provide software for reading the GPS timecode and performing the synchronization.) More details behind these requirements are outlined below in the section on synchronizing SPU data and video.

  To avoid the risks of integrating IFD and FRE software within a single processor, FREs must provide their own processor if they want to be integrated into the VSAM system.

- Radio Ethernet. The SPU will communicate with the OCU via radio Ethernet. We are currently planning to use the WaveLAN radio Ethernet line of products, with a WaveLAN ISA at the OCU and a WaveLAN WAVEPoint at each SPU. In cases where the OCU and SPU are in close proximity, a standard cable may be used instead. The IFD will make WAVEPoint boxes available to FREs at the demonstrations, subject to availability constraints.

- Video Transmitter. Video will be transmitted from the OCU to the SPU via a UHF link. We are currently planning to use a Dell-Star DSR 2400 FT/PL receiver at the OCU and a Dell-Star DSX 2423 transmitter at each OCU. These units have a nominal transmission range of 300 feet. Each SPU will require a seperate receiver unit at the OCU. In cases where the OCU and SPU are in close proximity, a standard coax cable will be used instead. The IFD will make transmitters available to FREs at the demonstrations, subject to availability constraints.

- Sensor Platform. The processor, if desired, may control parameters of the sensor platform, such as pan, tilt, and zoom. FREs desiring such control will need to provide their own sensor platform. This restriction has been imposed in order to alleviate the risks of jointly integrating IFD platform control software and FRE processing software to run on the processor.

As a specific example, the SPU that the IFD contractor is intending to assemble for its own use is shown in figure 6.

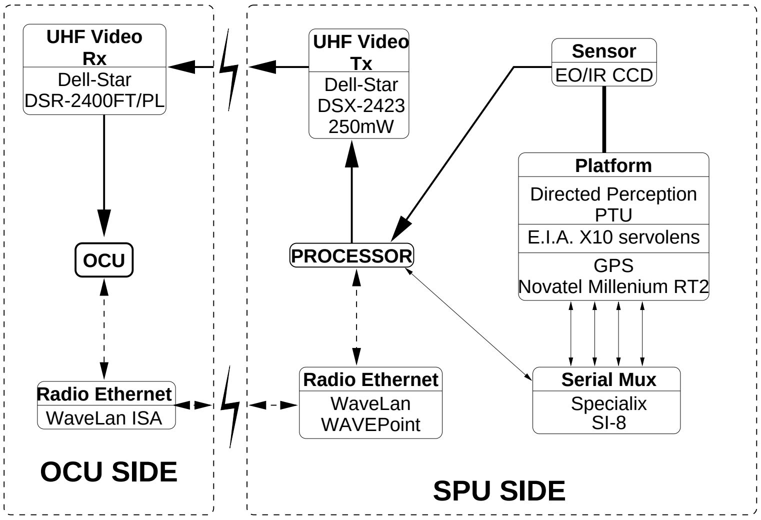


Figure 6: IFD Contractor SPU Configuration

4.3 Synchronization of data and video

The OCU must be able to combine data from multiple SPUs which may be processing video at different rates. Therefore, data from each SPU must be associated with a timestamp, synchronized to a system-wide clock, that represents the time at which the processed frame was digitized. To obtain this system-wide clock, we will initialize each processor to the universal GPS time, obtained using a GPS receiver that is temporarily plugged into the SPU serial port at initialization time. Each SPU will extrapolate from the initialized GPS time to obtain a time with which to tag its data.

A second problem is that that the SPU processors will require several frame times to process each frame they digitize. Therefore, the data arriving at the OCU will not be synchronized to the current frame being produced by the sensor. This will create problems if the OCU attempts to overlay the data on top of the video returned from the sensor. Our solution to this problem is to have each SPU send back on the video link the last frame it has processed, rather than the most recent frame produced by the sensor. This means that each SPU processor must have a video output.

5 Multi-sensor coordination and calibration

To simplify system design and modularity, all analysis and decision-making involving multiple sensors will take place on the OCU. Each SPU will know only of its own sensor and will not know about other SPUs. Each SPU will receive tasks and expectations from the OCU that enable it to obtain data that can be merged into a cohesive whole by the OCU.

To enable data from each SPU to be unified, the sensors must be calibrated with respect to each other and to a scene model representation such as a mosaic, orthomosaic, map, DEM, or combination thereof. We plan to experiment with two primary approaches to such calibration. The first will represent the 3D sensor positions and orientations relative to the model. The second will decompose the problem into two types of mappings that are based on fundamental matrices: camera-to-camera mappings that map one camera's coordinates to epipolar lines in another camera's view, and camera-to-model mappings that map each camera's view to an aerial mosaic or map. The fundamental matrices will be estimated on the basis of manually-entered point correspondences.

6 Data structures and communication

The tables below outline the types of data that will be present in the system's data structures and communications protocols. The exact form of the information is yet to be determined.

6.1 Sensor data structures

OCU	SPU
identifier (SPU number)	identifier
modality (EO, IR, etc.)	NA
task (idle, searching, tracking)	task
target search type (friend/foe, etc.)	NA
position - (x, y, z)	GPS lat/long/z
pose - (ρ, θ, ϕ)	pan/tilt + base inclination (stationary)
	pan/tilt + pitch/yaw/roll (airborne)
zoom/focus	zoom/focus
view volume	view plane
'current target' pointers	'current target' pointers
footprint	NA
potential FOV	NA

6.2 Target data structures

Scene level	Image level
identifier	identifier
type	type
'seen by SPU' pointers	NA
'last view' SPU pointer	NA
'last view' time	'last view' time
'last view' image	'last view' image template (32x32)
pointer to data base	image template (32x32)
3D position (x, y, z)	2D position (u, v)
3D pose (ρ, θ, ϕ)	NA
3D trajectory (v_x, v_y, v_z)	2D trajectory (v_u, v_v)

6.3 Communication packets

mode	OCU $\rightarrow$ SPU	SPU $\rightarrow$ OCU
idle	- idle command	NONE
search	- search command - region pan/tilt "A" to pan/tilt "B"	- target found response - image level SPU packet - image level target packet
track	- track command - image level target packet	- tracking response - image level SPU packet - image level target packet

7 Work Schedule

Figure 7 shows the detailed work scheduled proposed by the IFD team for Demo I.

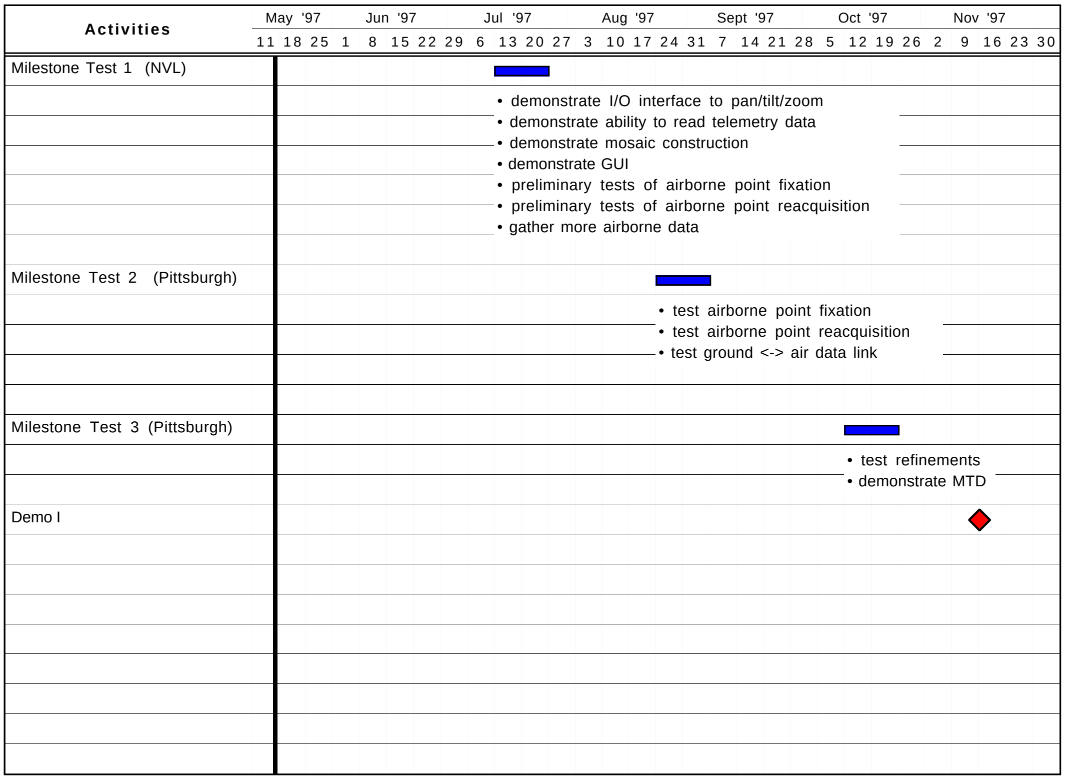


Figure 7: IFD work schedule

- END OF DOCUMENT -