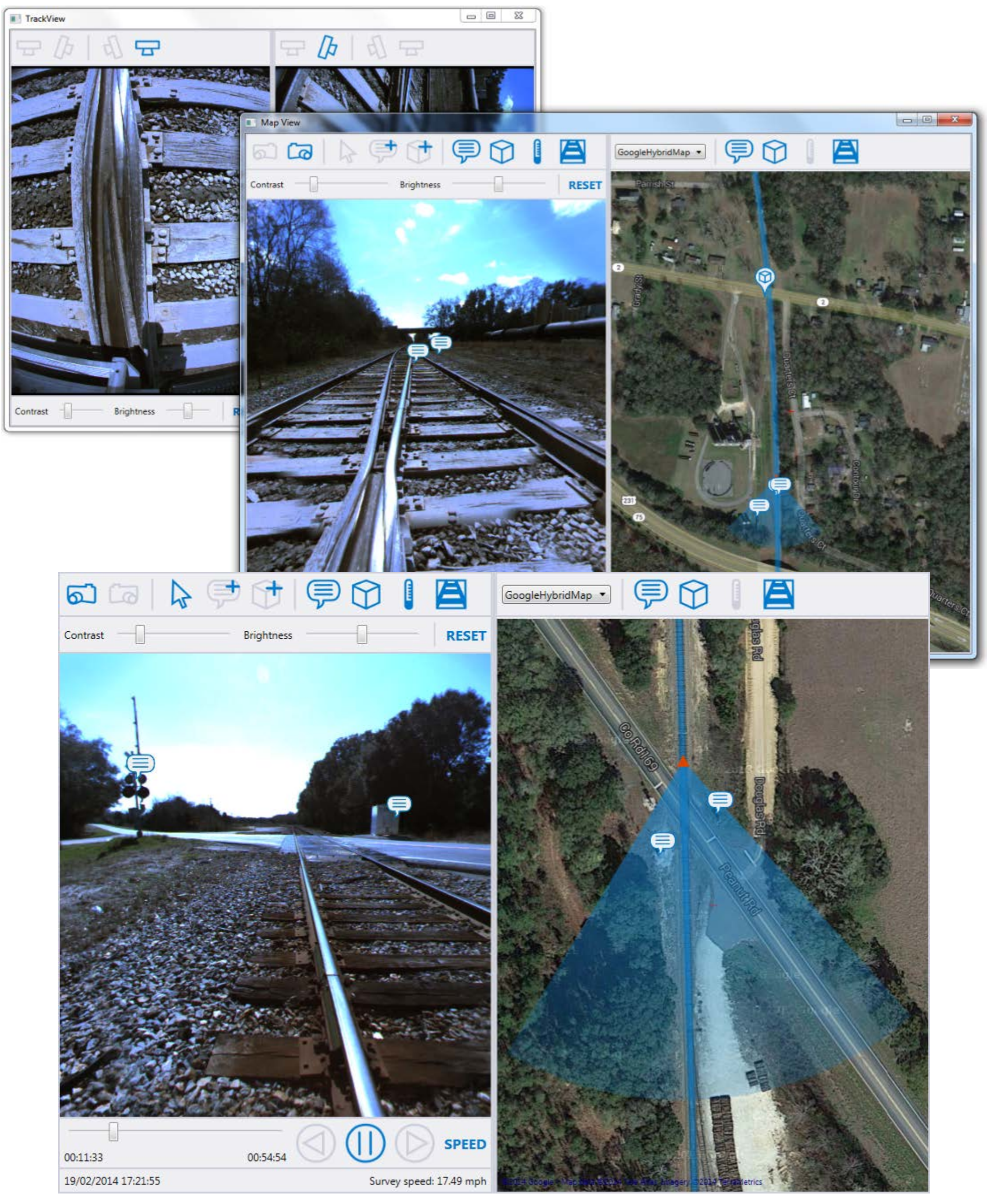
U.S. Department of Transportation

Federal Railroad Administration

Track and Track-Side Video Survey Technology Development

Office of Research
and Development
Washington, DC 20590



DOT/FRA/ORD-15/15

Final Report
May 2015

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	May 2015	Technical Report – August 2014

4. TITLE AND SUBTITLE
Track and Track-Side Video Survey Technology Development

5. FUNDING NUMBERS

6. AUTHOR(S)
Matt Mellor, Richard McCormack, Cameron Stuart (FRA)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
HiDef/Createc
Unit 4, Derwent Mills Commercial park
Cockemouth
Cumbria, CA13 0HT

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Department of Transportation
Federal Railroad Administration
Office of Railroad Policy and Development
Office of Research and Development
Washington, DC 20590

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
DOT/FRA/ORD-15/15

11. SUPPLEMENTARY NOTES
COR: Cameron Stuart

12a. DISTRIBUTION/AVAILABILITY STATEMENT
This document is available to the public through the FRA Web site at http://www.fra.dot.gov.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
Researchers at HiDef/Createc have completed prototype development and testing of a novel track video surveying technology called Track and Track-Side Video Survey (TTVS). TTVS is designed to capture clear video images of the track and track side areas from a moving platform at speed. This video data can be used to augment manual visual surveys and to add contextual value to existing data sets. The system has a user-friendly interface that allows for efficient data display and manipulation. Researchers completed trials of the prototype TTVS system on railroads in Maryland and Florida and demonstrated the system's ability to gather detailed imagery of the track and track-side at a resolution of 1 mm. Clear images were obtained at speeds up to 28 mph under a variety of weather conditions. Researchers also demonstrated the linking of external data to the video using geo-referencing techniques and used a novel algorithm to geo-track images within the field of view. With additional development, the TTVS system can be employed in various track inspection and documentation tasks. Recommendations for future development include improving video resolution at higher speeds, dynamic linking of external data, and creating algorithms to make measurements directly from the video images.

14. SUBJECT TERMS
Railroad, track, video survey, Egomotion, Global Positioning System

15. NUMBER OF PAGES
xxx

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	

NSN 7540-01-280-5500 S

METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC

LENGTH (APPROXIMATE)

1 inch (in)	=	2.5 centimeters (cm)
1 foot (ft)	=	30 centimeters (cm)
1 yard (yd)	=	0.9 meter (m)
1 mile (mi)	=	1.6 kilometers (km)

AREA (APPROXIMATE)

1 square inch (sq in, in^2)	=	6.5 square centimeters (cm^2)
1 square foot (sq ft, ft^2)	=	0.09 square meter (m^2)
1 square yard (sq yd, yd^2)	=	0.8 square meter (m^2)
1 square mile (sq mi, mi^2)	=	2.6 square kilometers (km^2)
1 acre = 0.4 hectare (he)	=	4,000 square meters (m^2)

MASS - WEIGHT (APPROXIMATE)

1 ounce (oz)	=	28 grams (gm)
1 pound (lb)	=	0.45 kilogram (kg)
1 short ton = 2,000 pounds (lb)	=	0.9 tonne (t)

VOLUME (APPROXIMATE)

1 teaspoon (tsp)	=	5 milliliters (ml)
1 tablespoon (tbsp)	=	15 milliliters (ml)
1 fluid ounce (fl oz)	=	30 milliliters (ml)
1 cup (c)	=	0.24 liter (l)
1 pint (pt)	=	0.47 liter (l)
1 quart (qt)	=	0.96 liter (l)
1 gallon (gal)	=	3.8 liters (l)
1 cubic foot (cu ft, ft^3)	=	0.03 cubic meter (m^3)
1 cubic yard (cu yd, yd^3)	=	0.76 cubic meter (m^3)

TEMPERATURE (EXACT)

[(x-32)(5/9)] °F = y °C

METRIC TO ENGLISH

LENGTH (APPROXIMATE)

1 millimeter (mm)	=	0.04 inch (in)
1 centimeter (cm)	=	0.4 inch (in)
1 meter (m)	=	3.3 feet (ft)
1 meter (m)	=	1.1 yards (yd)
1 kilometer (km)	=	0.6 mile (mi)

AREA (APPROXIMATE)

1 square centimeter (cm^2)	=	0.16 square inch (sq in, in^2)
1 square meter (m^2)	=	1.2 square yards (sq yd, yd^2)
1 square kilometer (km^2)	=	0.4 square mile (sq mi, mi^2)
10,000 square meters (m^2)	=	1 hectare (ha) = 2.5 acres

MASS - WEIGHT (APPROXIMATE)

1 gram (gm)	=	0.036 ounce (oz)
1 kilogram (kg)	=	2.2 pounds (lb)
1 tonne (t)	=	1,000 kilograms (kg)
	=	1.1 short tons

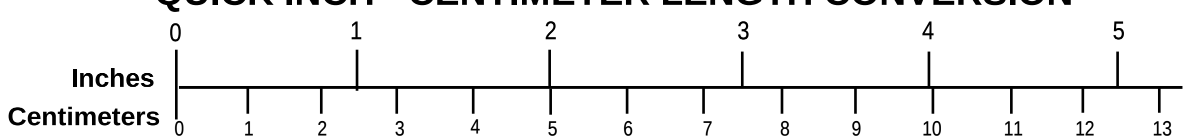
VOLUME (APPROXIMATE)

1 milliliter (ml)	=	0.03 fluid ounce (fl oz)
1 liter (l)	=	2.1 pints (pt)
1 liter (l)	=	1.06 quarts (qt)
1 liter (l)	=	0.26 gallon (gal)
1 cubic meter (m^3)	=	36 cubic feet (cu ft, ft^3)
1 cubic meter (m^3)	=	1.3 cubic yards (cu yd, yd^3)

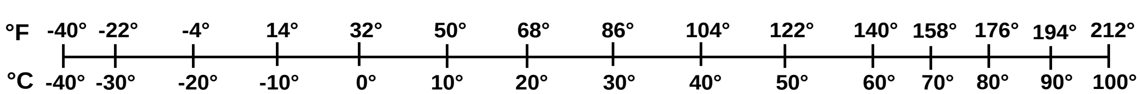
TEMPERATURE (EXACT)

[(9/5) y + 32] °C = x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSION



For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 6/17/98

Contents

Illustrations

Executive Summary

Track maintenance survey processes currently require a significant amount of manual observation, causing long and costly disruptions to normal track operation and exposing maintenance engineers to health and safety risks. The track and track-side video survey (TTVS) system detailed in this report can reduce the time spent on track by capturing continuous high resolution video of the track and surrounding areas, which provides a complete suite of pre-processed visual data that enables the user to conduct a virtual manual survey in the safety of their own office. The TTVS system will eventually enable operators to undertake surveys at track speed, enhance survey performance with new data, repeat surveys retrospectively, and view the history of an object or site by comparing new and old data.

The TTVS includes a graphical user interface (GUI) to provide the user with convenient access to a large quantity of high-quality images that closely replicate the on-track survey experience and provides new benefits, such as the ability to step back in time, or instantly leap ahead up the track. The interface demonstrates that an immersive video environment is effective as a tool for visual surveys of the track and track-side, and it can serve as a data navigation framework that the user visualizes and interacts with to interrogate geographic information system (GIS) data. All survey data is accurately geo-located, which allows users to locate a particular region in the video with a single click on a map instead of searching through the video.

Researchers completed trials of the prototype TTVS system on railroads in Maryland and Florida and demonstrated the system’s capability to gather detailed imagery of the track and track-side at a resolution of 1 mm. Clear images were obtained at speeds up to 28 mph under a variety of weather conditions. Researchers also demonstrated the capability to link external data to the video using geo-referencing techniques and used a novel algorithm to geo-track images within the field of view. With additional development, the TTVS system can be employed for a variety of track inspection and documentation tasks. The benefits of using this system over traditional methods of inspection includes reduced survey costs and lower health and safety risk associated with track maintenance activities. Recommended features for future development include improving video resolution at higher speeds, dynamic linking of external data, and creating algorithms to make measurements directly from the video images.

1. Introduction

This report describes the development and testing of a prototype track and track-side video survey (TTVS) system for use in railroad surveys and inspections. This project was funded by the Federal Railroad Administration's Office of Research and Development and performed by HiDef Aerial Surveying, Inc. (http://hidefsurveying.com/about-us-usa) in partnership with Createc Technologies (http://www.createc.co.uk). Additional support for this project was provided by ENSCO, Inc. (http://www.ensco.com/). ENSCO provided design, integration, and test execution support for testing activities conducted on the Federal Railroad Administration (FRA) R4 Hy-Rail research vehicle. The project duration was 1 year (2013- 2014).

1.1 Background

Track maintenance survey processes require a significant amount of manual observation, causing long and costly disruptions to normal track operation and exposing maintenance engineers to associated health and safety risks. Imaging systems enable the user to undertake a virtual manual survey by capturing continuous high resolution video of the track and surrounding areas, and providing a complete suite of pre-processed visual data. Key requirements of these imaging systems include 1) high resolution cameras for detailed inspection of ties and fasteners, 2) geo-location of image data to permit comparisons with prior surveys and to link other data types to the image files, 3) track speed data capture rates, and 4) an easy to use data interface.

1.2 Objectives

The project objective was to develop and demonstrate a track video imaging system that provides sufficient resolution to perform track inspection activities. The track will be imaged at 1 mm resolution, which will enable the user to identify subtle defects. The area surrounding the track will be imaged at a lower resolution.

The system is designed for installation on a track vehicle to capture continuous high resolution video and provide a complete suite of pre-processed visual data to the user. Demonstration testing was conducted using Hy-Rail vehicles.

Specific goals for this project were as follows:

- Design and integrate the spherical camera system and vehicle systems. FRA's R4 research Hy-Rail vehicle was selected as the test platform for on-track tests in the United States. Preliminary on-vehicle testing was conducted in England using a road truck.

- Develop algorithms that process images from the spherical camera and geo-locates objects within the field of view.

- Develop algorithms to allow users to select and track objects within the video environment.

- Develop graphical user interface (GUI) that is user-friendly.

- Conduct demonstration tests of the system, which includes linking external data sources to the video stream

1.3 Technical Approach

The technical approach was structured with four primary milestones:

<u>Milestone 1</u>:

a. Design the camera system

b. Integrate system with a test vehicle

c. Conduct preliminary field testing and data gathering activities

d. Develop hardware interfacing equipment for R4 field testing

<u>Milestone 2</u>: Complete initial software development, including Egomotion and object tracking algorithms.

<u>Milestone 3</u>:

a. Final software development

b. Develop a GUI

c. Implement object capture functionality

d. Incorporate geographic information system (GIS) overlay in video system

e. Complete software integration testing

<u>Milestone 4</u>:

a. Demonstration of prototype system on Western Maryland Scenic Railway (WMSR)

b. Integration of external data into the video system.

c. Final report and presentation

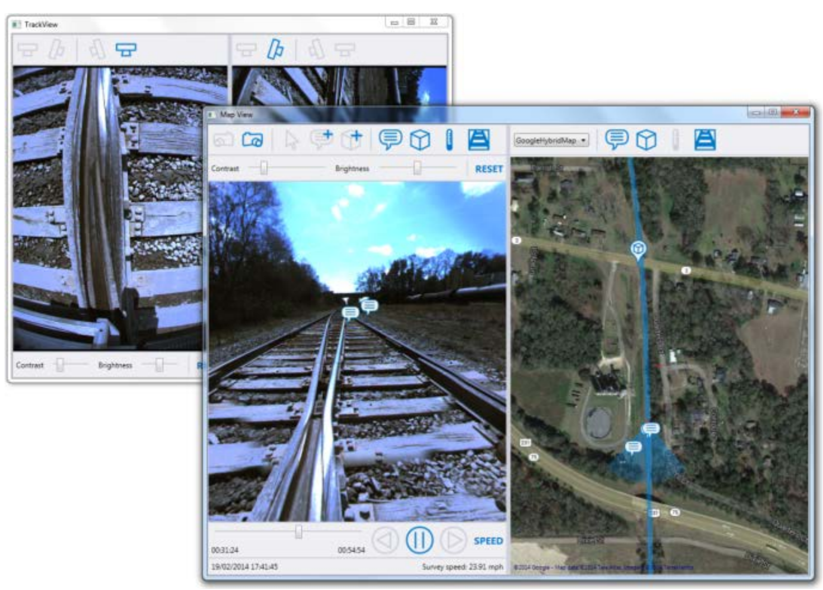


Figure 1. TTVS User Interface

2. Project Execution

2.1 Milestone 1: Design and Preliminary Testing

Although similar camera systems have been successfully deployed on trains, they have not been used to generate high resolution images of nearby objects. The combined effects of fast platform motion, vibration induced camera shake and short camera exposures combined to reduce the image quality and introduce technical risk. A core objective of this project was to demonstrate adequate imaging quality with a spatial resolution of 1mm on the track, which made image quality a potential risk.

In the first phase of the project, baseline data was gathered to demonstrate the basic capabilities of the camera system and to collect data for algorithm and software development in the subsequent phases of the project. At this stage, the scope of work was:

- Design a camera mounting system for the R-4 Hy-Rail vehicle (with support from ENSCO)
- Design adapters to mount the system on a suitable road vehicle
- Integrate the mechanical and sensor system prototypes
- Gather example rail survey data to demonstrate system performance.

The output of this milestone was:

- A prototype camera mounting system
- A set of development data
- Feedback on modifications to the camera mounting system that might be required before the final demonstration is held.

2.1.1 Camera System

The core instrument in the TTVS system is a carefully calibrated array of synchronized wide-angle cameras that captures full-frame video footage. As raw data comes from the cameras, software corrects and stitches the individual images to form a single spherical image. Figure 3 shows an example of spherical imaging.

The TTVS system incorporates a pair of the spherical camera systems, which are configured to yield an unobstructed view of the track and its surroundings. The cameras are mounted onto a horizontal beam and attached to the survey vehicle via anti-vibration mounts. The stereo cameras provide detailed views of both sides of the track and the system can derive range information from the image data.

Figure 2. Spherical video camera (left). Stereo spherical cameras attached to mounting beam (right).

Figure 3. A spherical image shown as a super-wide 'equirectangular' image.

The spherical video cameras are located approximately 1 m above the track and provide a close and detailed view of the rails, crossties and ballast, which enables the user to inspect the fine detail of the track. The use of commercial off the shelf (COTS) spherical cameras to provide this level of detail is the limiting factor on survey speed. Survey speed is dictated by camera resolution and frame rate, which imposed a theoretical limit of 30 km/hr for the purposes of this project.

Both cameras have an extremely wide field of view, which allows everything in front, to the side, or above the camera to be seen. A partial view of objects behind the survey system is also provided.

2.1.2 Preliminary Testing

Researchers used the vehicle/camera system to gather trial data for development of processing algorithms and software. The data was split into two sets: one for software development and the other for software testing.

The camera and recording systems were integrated into the survey vehicle (Figure 4). The cameras are rear-mounted on a beam at 1 m height above the ground; they yield a 1 mm resolution or better image of the area directly beneath each camera and a resolution of approximately 1.5 mm for the diagonally opposite surface.

Figure 4. Camera system mounted on the rear of Createc's technology demonstration vehicle.

Figure 5 shows the spherical image captured during the rail test. The rails are actually straight, but appear curved in the image due to the spherical-to-2D projection process.

Figure 5. A spherical image showing the view down a rail track.

The results from these initial trials were positive and researchers had a good indication of the image quality of the system. The system was exceeding the 1 mm resolution target and capturing fine details such as wood grain.

2.2 Milestone 2: Initial Software Development

In this task, researchers developed algorithms and software which allows the user to select and track objects as they come into view then, if required, zoom into the object. Effectively, the user gains the ability to rotate a virtual object about its center.

If the system is to serve as an environment for navigating spatial data, it must be able to accurately estimate the pose and position of the cameras in each frame. A process known as Egomotion, which calculates the motion of the camera in space directly from the images, accomplishes this task. For the project to be successful, Egomotion must be demonstrated with sufficient accuracy for GIS data to be converted into an accurate video overlay.

To enable GIS data to be created within the video it is necessary to be able to track objects across several frames so that a 3D location for the object can be inferred. This tracking process should be sufficiently accurate and robust to allow for single-click object identification.

2.2.1 GPS and Egomotion

Video footage is geo-referenced using a combination of GPS and Egomotion. Although Egomotion is less accurate than GPS over long distances, it is much more accurate for small movements and, unlike GPS, measures rotations accurately.

The fusion of GPS and Egomotion negates the need for an expensive inertial navigation system (INS) or differential-GPS (Differential Global Positioning System) system. Egomotion also provides the TTVS system with the ability to derive INS-like behavior by tracking objects in the video. Therefore, the system calculates the three dimensional location of the camera at each moment, locating all objects and structures visible in the footage. This enables automatic geo-referencing and integration with GIS. By providing an accurately geo-located sensor package, the TTVS system provides a platform for adding further sensors such as LIDAR (Light Detection and Ranging) or thermal imagery.

To measure camera motion directly from the spherical image data, Createc's Egomotion algorithms detect and track a large number of distinct features in the surrounding environment. By matching these features between the stereo camera pair, it is possible to calculate feature locations within a three dimensional space. Tracking these features and their positions over time then provides an accurate measurement of camera pose and location. Figure 6 illustrates the internal workings of the Egomotion algorithms. The images on the left of the figure show views from the two spherical cameras, superimposed with colored dots representing the features tracked in the scene. The short green lines illustrate the motion of these features between the previous and current frames; this motion is used to track the camera in three dimensions. The image on the right of the figure depicts a plan view of the path followed by the cameras (in green) and the locations of tracked features.

The TTVS Egomotion algorithms were developed using data from a survey of a light industrial and residential environment and a limited amount of rail data. The survey deliberately attempted to confuse the system with various "figure of eight" maneuvers. Despite the challenging data, the algorithms successfully measured camera motion and provided robust camera pose estimation. The peak test speed of 32kph, just over design speed, was comfortably within system capabilities. Over the length of a half mile survey route, small positioning errors resulted in a compound error of approximately twenty five feet. This integral error is inherent in any dead reckoning positioning system and a gross error of twenty five feet in almost four minutes suggests a very low instantaneous error rate. Refining the Egomotion algorithms and incorporating frequent Global Positioning System (GPS) measurements result in an accurate camera positioning system with almost no overall integral errors.

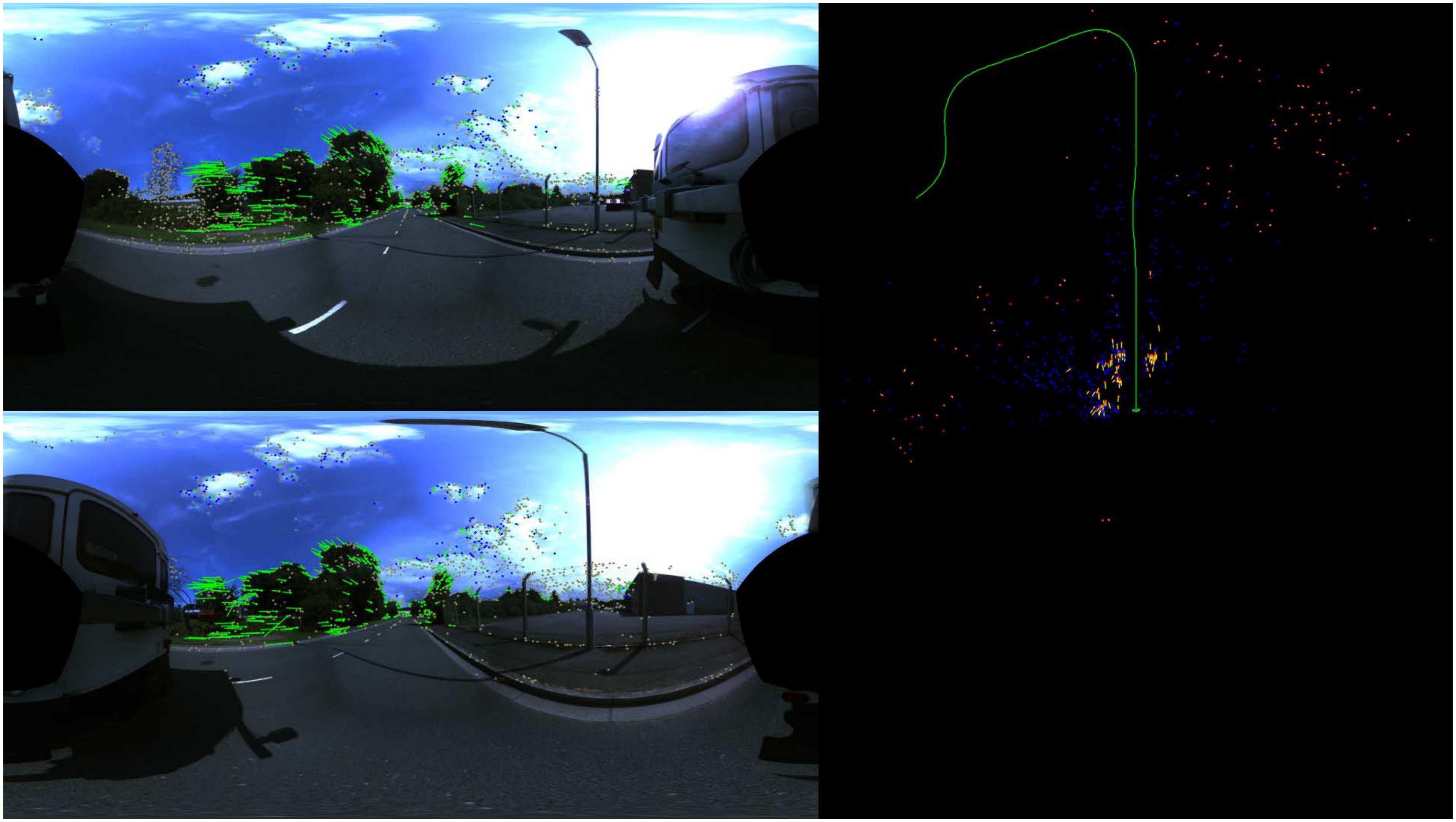


Figure 6. Internal workings of the Egomotion algorithm.

To interface with a GIS system, the relative locations from the Egomotion must be given absolute coordinates which can be related to the GIS mapping. Createc developed algorithms to accomplish this by fusing sparse GPS readings (left image in Figure 7) with output from the Egomotion process. This fusion produces a geo-location and pose for each frame in the video data. Incorporating the Egomotion data serves to interpolate between GPS readings while also removing a large proportion of the GPS noise (right image in Figure 7).

Figure 1. GPS-Egomotion fusion.

2.2.2 Object Tracking Algorithms

Createc developed algorithms which allow objects to be tracked across different frames and views. This capability allows the user to select an object within a view. Once an object is selected, the system will automatically locate it in all frames in which it is visible, thus allowing the visual impression of rotating the object about its center.

The object tracking uses feature-based tracking algorithms to capture all possible views of the object; the selected object is sampled and its salient features are tracked both forwards and backwards in time. In Figure 8, an object is selected from within a rectified window of a spherical image by drawing around it. The algorithm tracks the salient features of this selection (red dots in left image) to the next frame in the video and estimates the physical boundaries of the tracked object.

Figure 8. Object tracking algorithm.

Figure 8 is an example of object tracking output from the TTVS.

Figure 9. Object tracking views.

2.2.3 Simulated Pan, Tilt, and Zoom

The spherical video cameras capture images that cover over 80% of the full sphere and produce panoramic images like that shown in Figure 3. The user does not view these panoramas but instead views a rectified region of the panoramic image, which creates the illusion of looking at a standard image; see Figure 10. The large window shows a panoramic image with a rectified region highlighted in white. The smaller window depicts the result of this rectification. By controlling the size and location of this region we effectively simulate a pan-tilt-zoom (PTZ) camera.

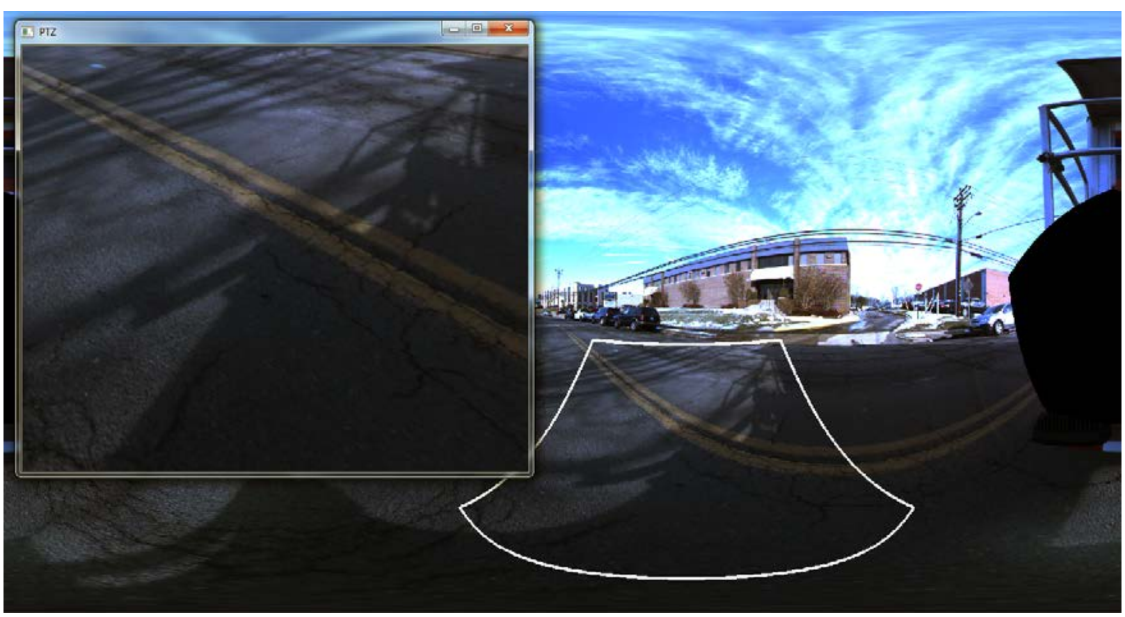


Figure 2. Simulated pan-tilt-zoom feature.

2.3 Milestone 3: Final Software Development

After the core algorithms were successfully demonstrated, development efforts were focused on the prototype TTVS user interface. A video survey consists of continuous recording with the spherical cameras, followed by post processing to produce the geo-referenced virtual environment. This interface provides the backbone to the data system.

The TTVS system is meant to provide the user with a natural and intuitive interface to a large quantity of high-quality imagery and replicates the experience of actually being there as closely as possible, while adding new benefits, such as the ability to step back in time or instantly leap ahead up the track. In this phase, the research goal was to develop an appropriate GIS front-end and user interface that presents the user with an intuitive way of visualizing, exploring and creating data.

Geo-referenced footage enables fusion of the images with a GIS system. A simulated aerial view can be generated to associate TTVS or other data with aerial imagery, allowing the user to navigate in a map environment.

GIS objects can be visualized as overlays within the virtual survey environment and created using simple video mark-up tools. Figure 11 depicts a screenshot from the TTVS user interface showing the video, map, and track view interfaces as well as GIS object overlays.

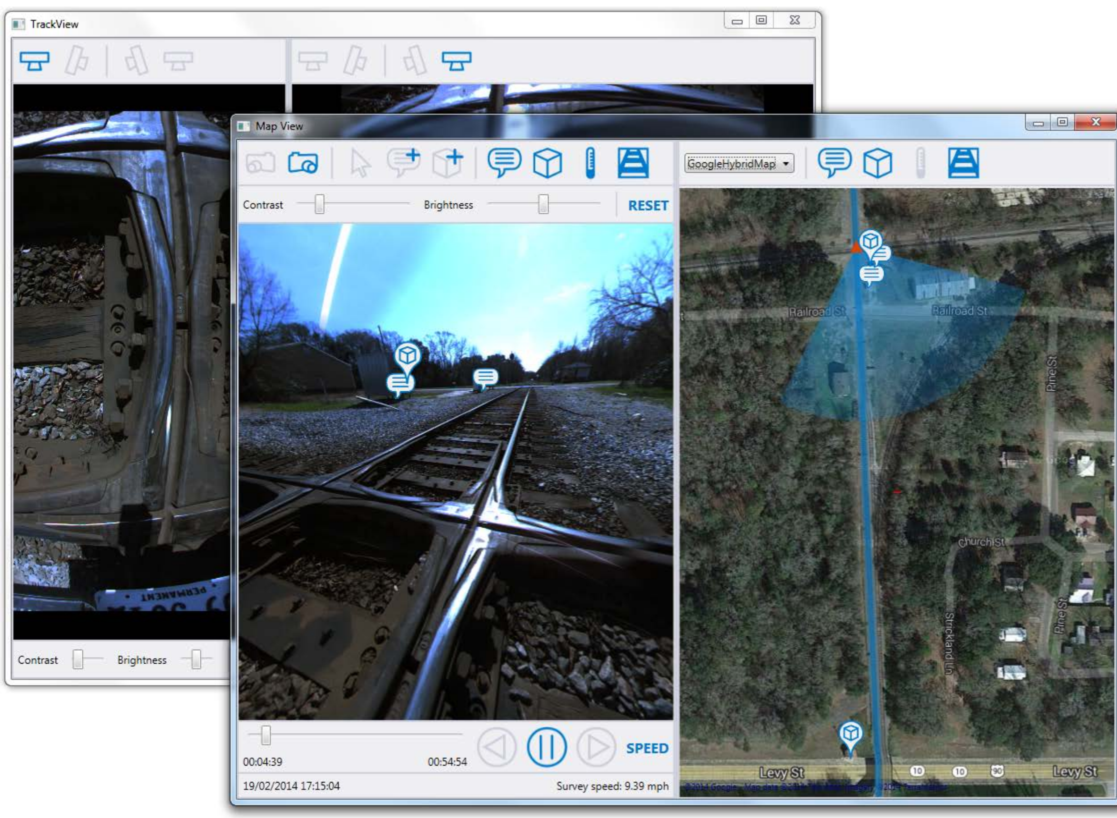


Figure 3. Screenshot from the TTVS user interface showing video, map, and track view interfaces.

2.3.1 Database Integration

This demonstration project used a local GIS-enabled database to store all survey related data. With further development, the GIS and TTVS systems can be integrated. Given the prototype nature of this project, Createc engineers chose to develop a self-contained database environment instead of large, integrated GIS environment.

The database contains the following tables:

Survey information - This table includes data on each survey such as name, date, and location as well as the location of any video files. Hardware parameters such as stereo baseline distance and camera height are also included.

Survey path - This table contains multiple outputs from the Egomotion process, i.e. data on the position and pose of the cameras for each frame in the video survey.

GIS objects - This table includes data on any GIS objects created in, or imported into, the system. Data includes parameters such as location, timestamp, name etc.

2.3.2 Prototype Graphical User Interface

The TTVS GUI demonstrates the effectiveness of an immersive video environment as a tool for visual surveys of the track and track-side and as a data navigation framework that visualizes, interacts with and creates GIS data. The user is given an interface to a large quantity of high-quality imagery that replicates the experience of actually being track-side as closely as possible, while adding new benefits, such as the ability to step back in time or instantly leap ahead up the track.

The concept GUI and accompanying toolset present an intuitive interface through which to examine survey data. Createc and FRA decided that the GUI would consist of three core graphical windows; track view, context view, and map view. The following sections describe the functionality of these windows.

2.3.3 Track View

The track view window is dedicated to providing a detailed view of the rails and immediate surrounding area; ideally this view will have a dedicated monitor. The window has the following core functionality:

- Two independent views of the tracks with pan and zoom functionality allowing the user to focus on, and zoom into, areas of interest. The contrast and brightness of each image is independently controllable.

- Buttons which enable the user to change their view of the tracks. For each sub-window, the user can choose between a top-down view of a rail or a view of the inside face of a rail. Figure 12 shows the top-down view of the left and right rails.

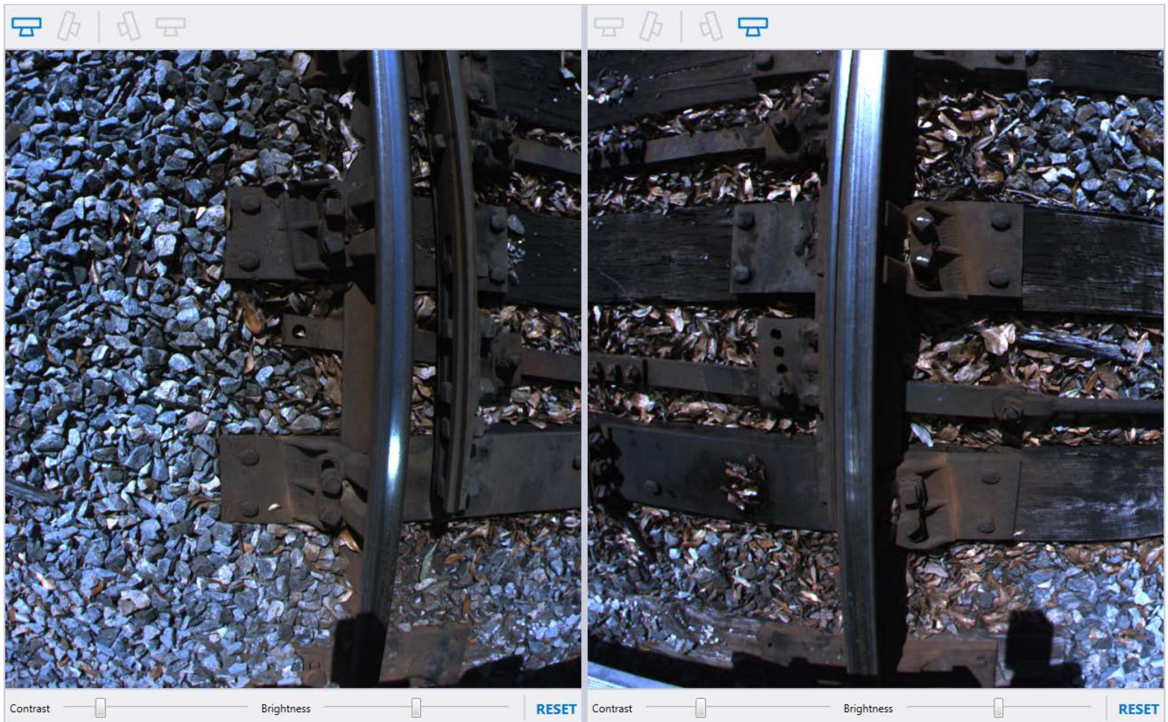


Figure 4. Track View window.

2.3.4 Context View

The context view provides the user with the capability to see the surrounding area and interact with video. This view has the following core functionality:

- A simulated PTZ camera which allows the user to focus on, and zoom into, areas of interest
- Video controls allowing survey playback, both forwards and backwards, at various speeds
- Controls allowing the user to switch between cameras
- Controls enabling the fine-tuning of image contrast and brightness
- The ability to attach notes to objects within view
- The ability to select objects for object tracking
- Frame information such as timestamp and survey speed
- Buttons that toggle the display of video overlays
- The ability to display, focus on, and edit tracked objects and survey notes. The user can adjust the location of GIS objects through a simple click-and-drag interface

One of the core tools in the context view is the selection tool, which enables object selection from within the PTZ view through a simple click-and-drag interface. This feature is illustrated in Figure 13. Once an object is selected, the user can attach a survey note to that object or track it through the video.

When the selection tool is activated, the software searches for the selected object in the partnering camera. The user is presented with the result of this search and asked to either confirm or correct the stereo match. The software uses this stereo match and the knowledge of the position and pose of the cameras at each frame in the video to automatically calculate geographic coordinates for the selected object.

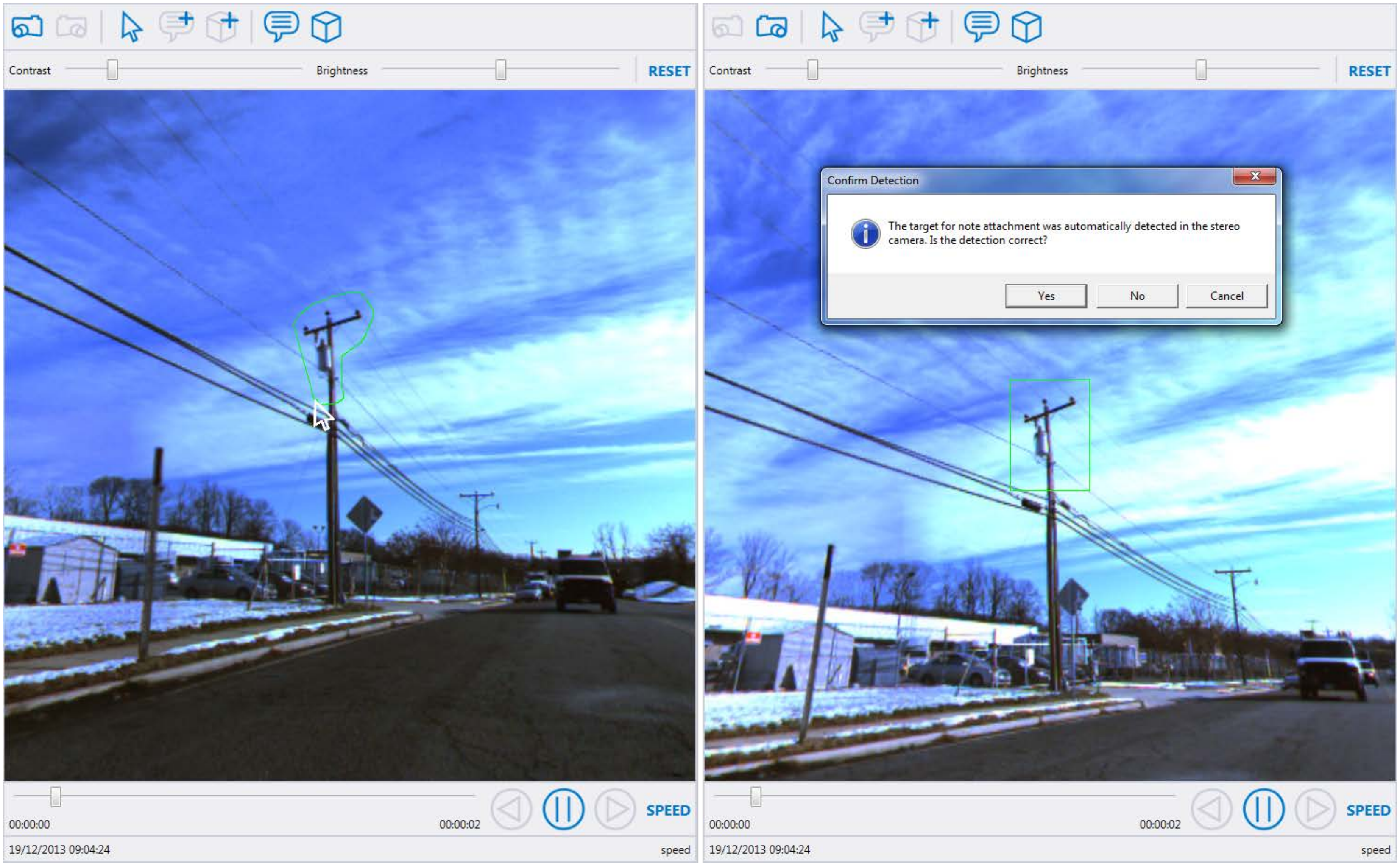


Figure 5. Context view with object selection.

2.3.4.1 Attaching a Survey Note

Once an object is geo-located, the user is presented with a pop-up window where he or she can create a survey note. This survey note is exported as a GIS object and added to the map interface.

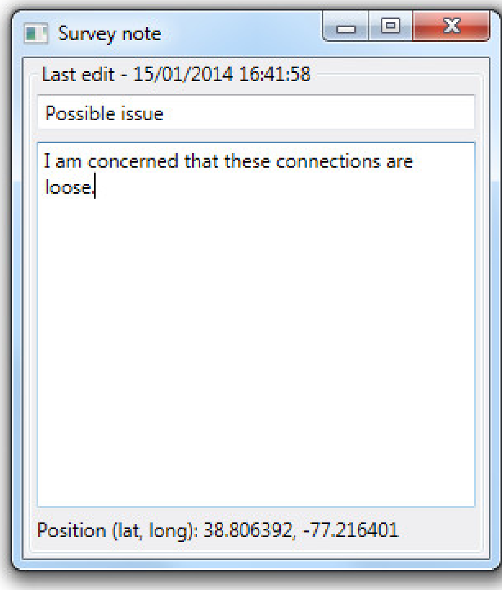


Figure 6. Survey note pop-up window.

2.3.4.2 Tracking an Object

Once an object is geo-located, the software searches adjacent frames in the video for the same object. The GUI presents the user with the object tracking window depicted in Figure 15. Using this window the user can scroll through different views of the tracked object. The user can adjust zoom and correct the tracking if desired. The tracked object is exported as a GIS object and added to the map interface.

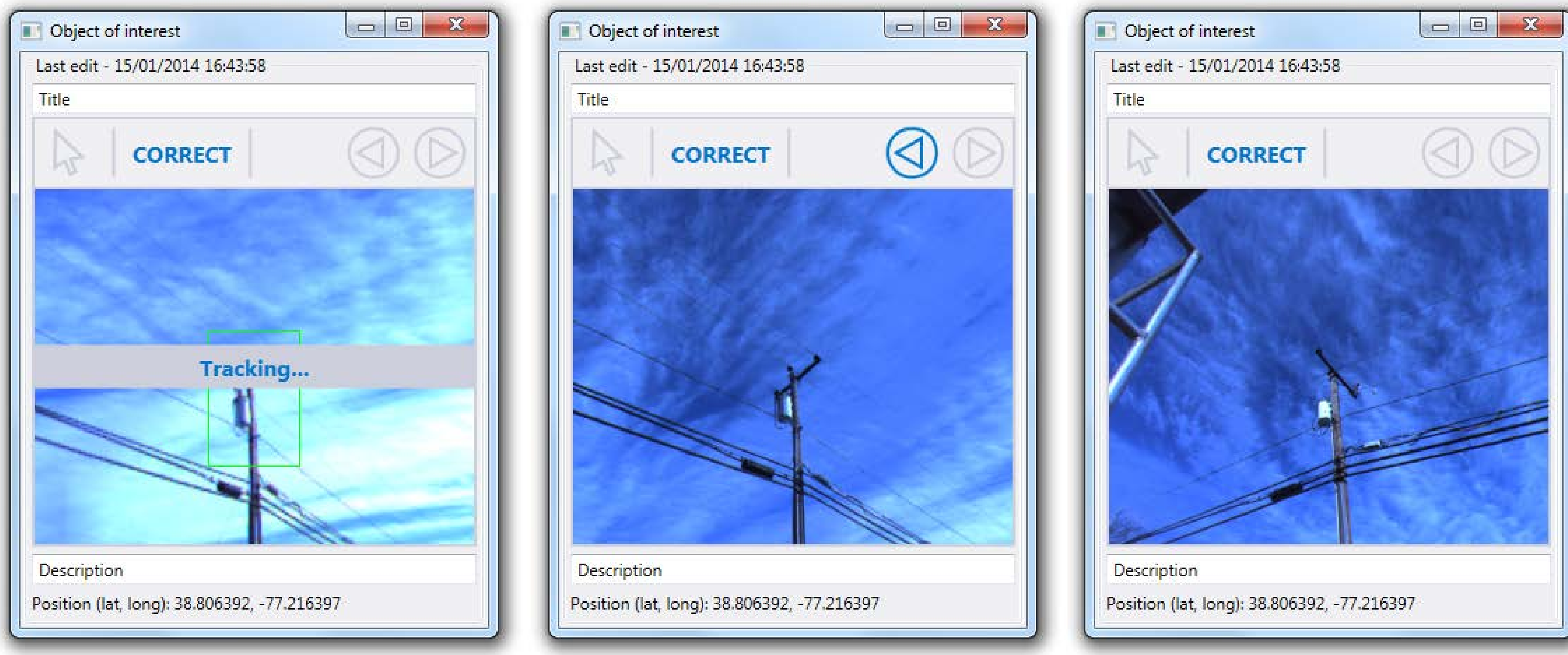


Figure 7. Tracked object pop-up window.

2.3.4.3 Video Overlays

The context view also includes the capability to display GIS objects as overlays on the PTZ view. The GUI displays all GIS objects at their geo-referenced locations, which are calculated from the software's knowledge of the camera's position and pose. The positioning of GIS overlays is updated in each frame and in every manipulation of the PTZ view, reinforcing the impression of viewing objects that are fixed in space.

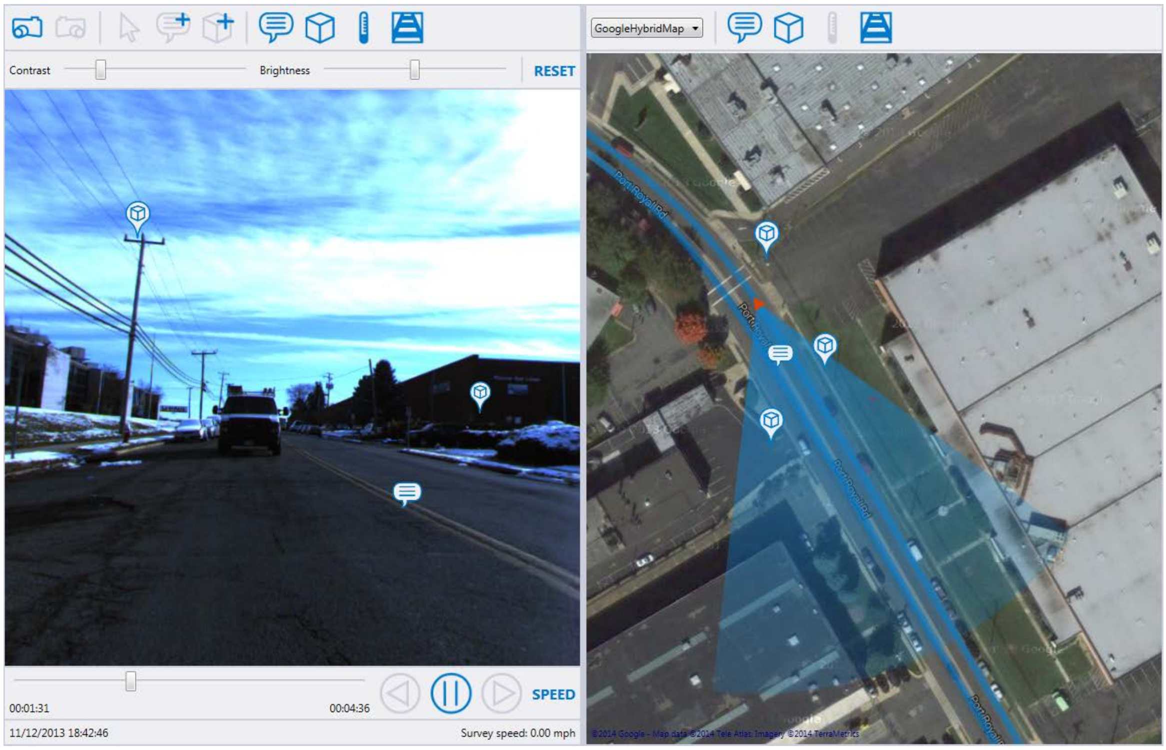


Figure 8. GIS objects in PTZ and geo-referenced views.

The user can interact with video overlays through a simple point-and-click interface. By clicking and dragging, the user can adjust the location of a GIS object and the map view updates with any changes. Additionally, the user can right-click a GIS overlay to skip the video to the nearest frame, panning and tilting the PTZ view to focus on the object. This function is particularly useful when a user is exploring the salient regions in a survey video or browsing previously created GIS objects.

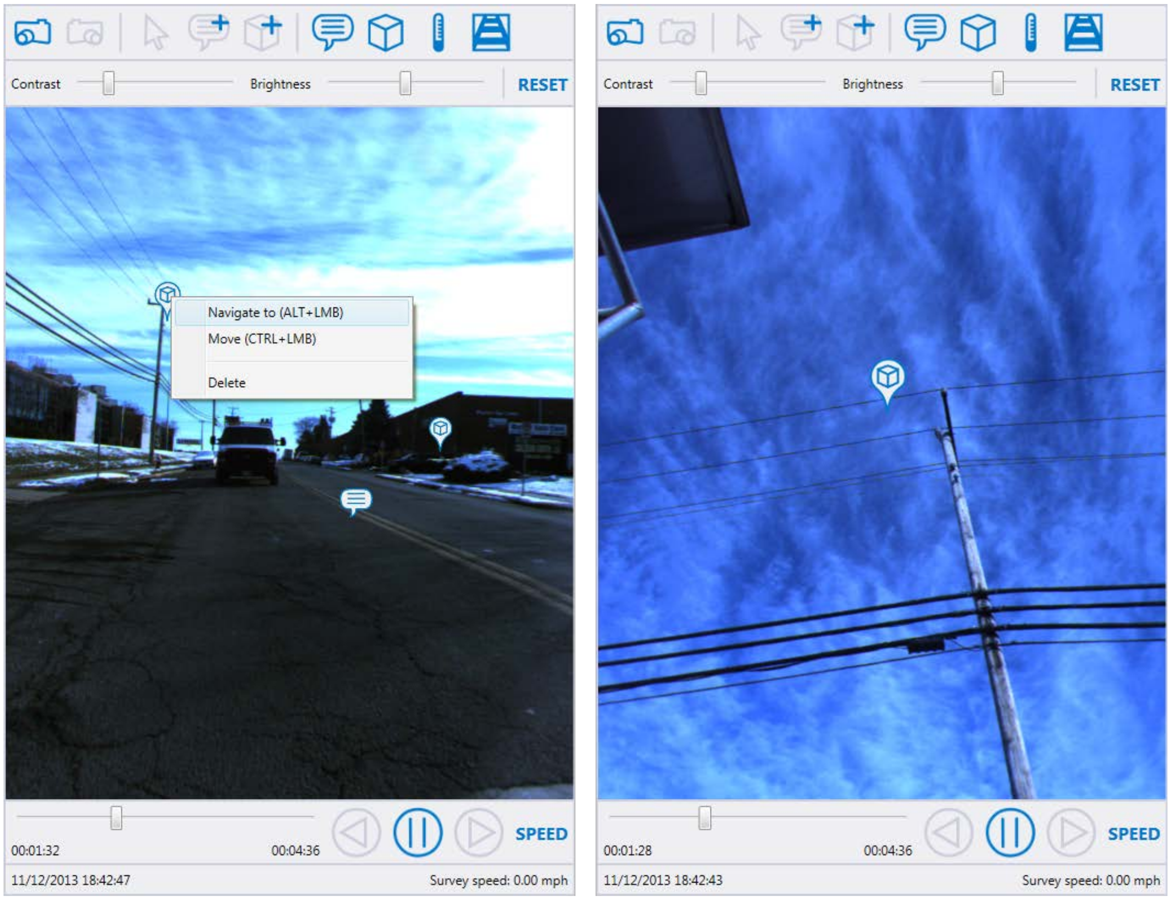


Figure 9. Selection and display of GIS object.

2.3.4.4 Map View

The map view provides the user with a global overview of the surveyed track including overlays that feature GIS objects. The window has the following core functionality:

- The ability to pan and zoom across the interactive map.
- Overlays that depict the survey route and the current location and field of view of the context camera.
- The ability to select a location on the map and skip to the nearest frame, panning and tilting the context view to focus on the selection. This function allows the user to skip to a particular region in the video with a single click of the mouse on the map instead of performing a laborious search through the video.
- The ability to display, focus on, and edit tracked objects and survey notes. The user can adjust the location of GIS objects through a simple click and drag interface.
- Buttons that toggle the display of map overlays.

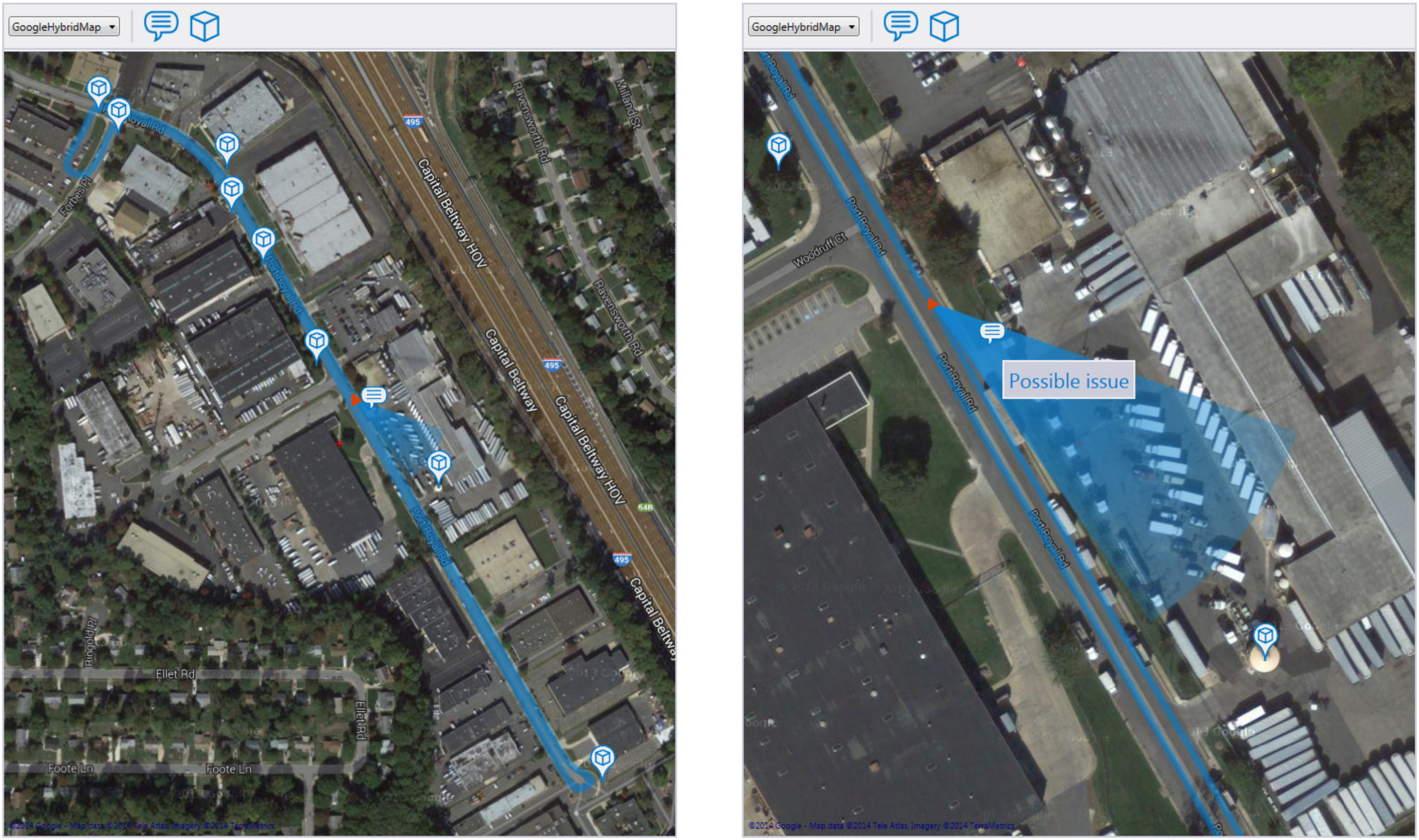


Figure 10. Map view with survey route and GIS overlays.

2.3.5 Process Workflow

For this demonstration project, the survey process was performed by Createc personnel familiar with the TTVS system and as such, no effort has been spent to automate the process or develop a simple user interface. Despite this, ENSCO personnel have successfully operated the system on a number of separate occasions. The process described below will be simplified, wrapped in an intuitive GUI, and automated as much as is possible as part of a commercialization development effort.

The survey process currently consists of the following steps:

Record track video - Before each survey, image quality and camera exposure settings are manually set using the camera’s proprietary software. A save location is specified within a configuration file and recording initiated on each camera in turn. On survey completion, the operator stops recording on each camera.

Create a new survey in the database - Prior to post processing, a new survey entry is added to the database. This entry includes parameters such as the name and date of the survey, the location of video files on the computer system, the stereo baseline distance, and the height of the cameras above the rails. Each survey has a unique ID.

Calibrate cameras - Calibration is only necessary if the cameras have been removed from the mounting beam since the last survey. The software presents the same frame from each camera and asks the operator to match points between the images, and then it uses these points to calculate the relative position and pose of the cameras and saves this to a calibration file.

Run Egomotion process - The operator passes the ID of the new survey to the Egomotion software, which then proceeds to process all related video. After the data has been processed, the software exports a text file with the Egomotion coordinates, camera pose, and absolute GPS coordinates for each frame in the video.

Fuse Egomotion and GPS coordinates - The Egomotion coordinates and the absolute GPS coordinates are passed through a series of filters to produce the final survey path. When the path is complete, the operator saves it to the database for use by the GUI.

Review survey data - The survey data is now ready for review by the virtual maintenance engineer.

3. System Demonstrations and Results

3.1 Western Maryland Scenic Railroad

In order to demonstrate the value of the TTVS system, the project team executed a full concept demonstration on a section of the Western Maryland Scenic Railroad (WMSR). The system mounted onto the FRA's R4 hi-rail vehicle. The integration and commissioning of the system on the R4 was a quick and simple process facilitated by ENSCO personnel. Figure 19 shows the TTVS system mounted onto the back of the R4.

3.1.1 Survey Execution

ENSCO undertook two surveys on the WMSR between December 2013 and January 2014. During each survey, operators conducted a number of runs over a section of track approximately 1.5 miles in length, driving the R4 both forwards and backwards. Survey speeds varied between 5mph and 15mph. Some minor GPS issues were encountered in early runs when the GPS struggled to gain a satellite lock under tree cover. This issue only affected a small proportion of runs.

Both surveys were conducted on very overcast days with cloud cover significantly reducing ambient lighting. The low light levels had the undesired effect of increasing image noise and introducing motion blur for some runs at higher survey speeds. For the most part, these issues were overcome by overriding the camera's automatic exposure settings. Future TTVS systems will be able to automatically calculate optimum exposure settings for various lighting conditions.

3.1.2 Example Imagery

Despite adverse weather conditions, the WMSR trial surveys were a success, producing high resolution track and track-side imagery. Imagery of the track is at a resolution of approximately 1mm, meeting design specifications. Low light conditions resulted in slightly noisy images but this did not affect system functionality.

Figures 20 and 21 demonstrate the system's ability to zoom into regions of images via the context view of the GUI. Figures 22 and 23 depict images from the four selectable views in the track view window of the GUI. Figures 24 and 25 show how the system can zoom down into track images and focus on an area of interest.

Figure 11. TTVS system mounted on the R4 hi-rail vehicle.

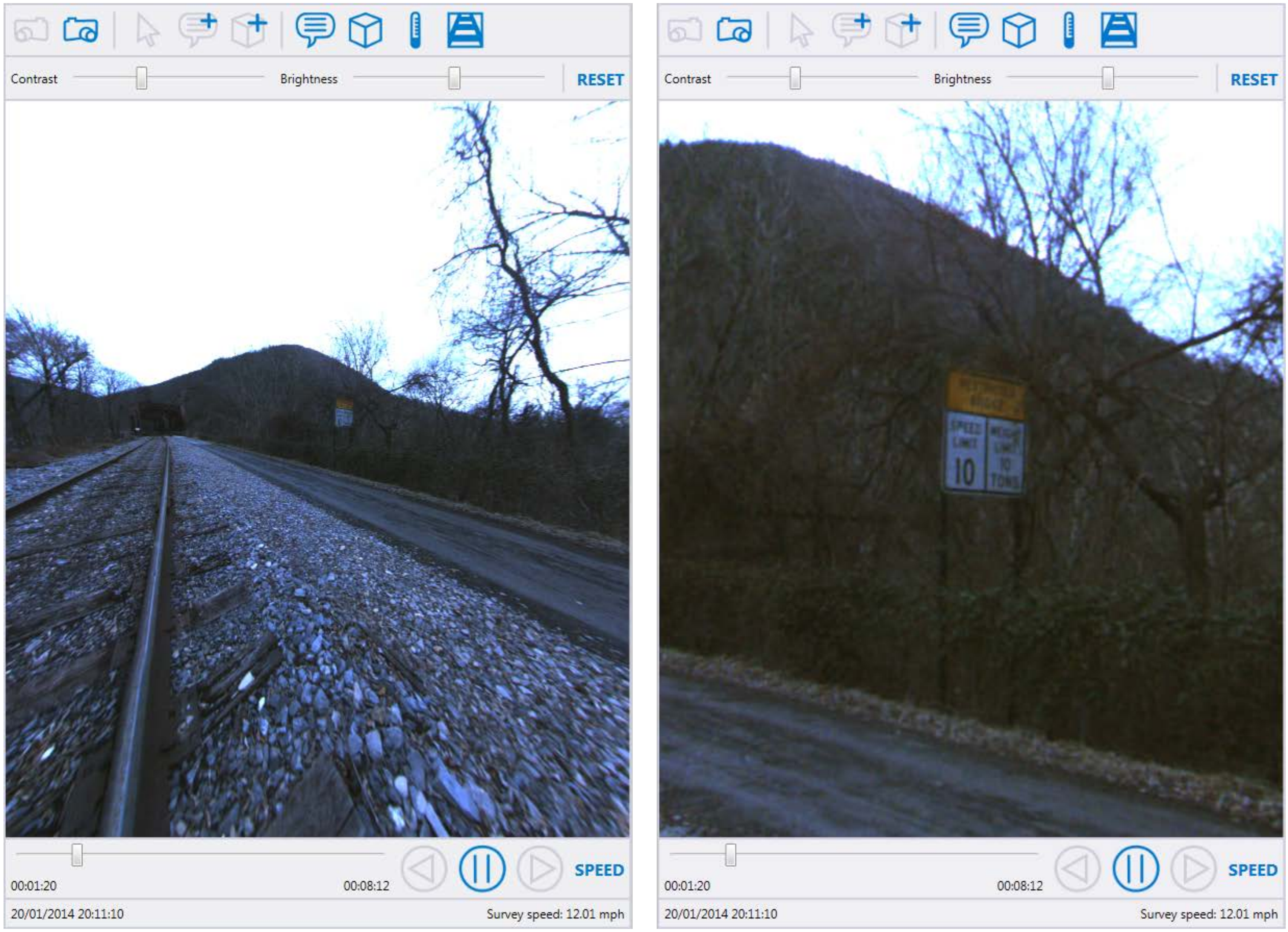


Figure 12. Context view screenshots of WMSR data.

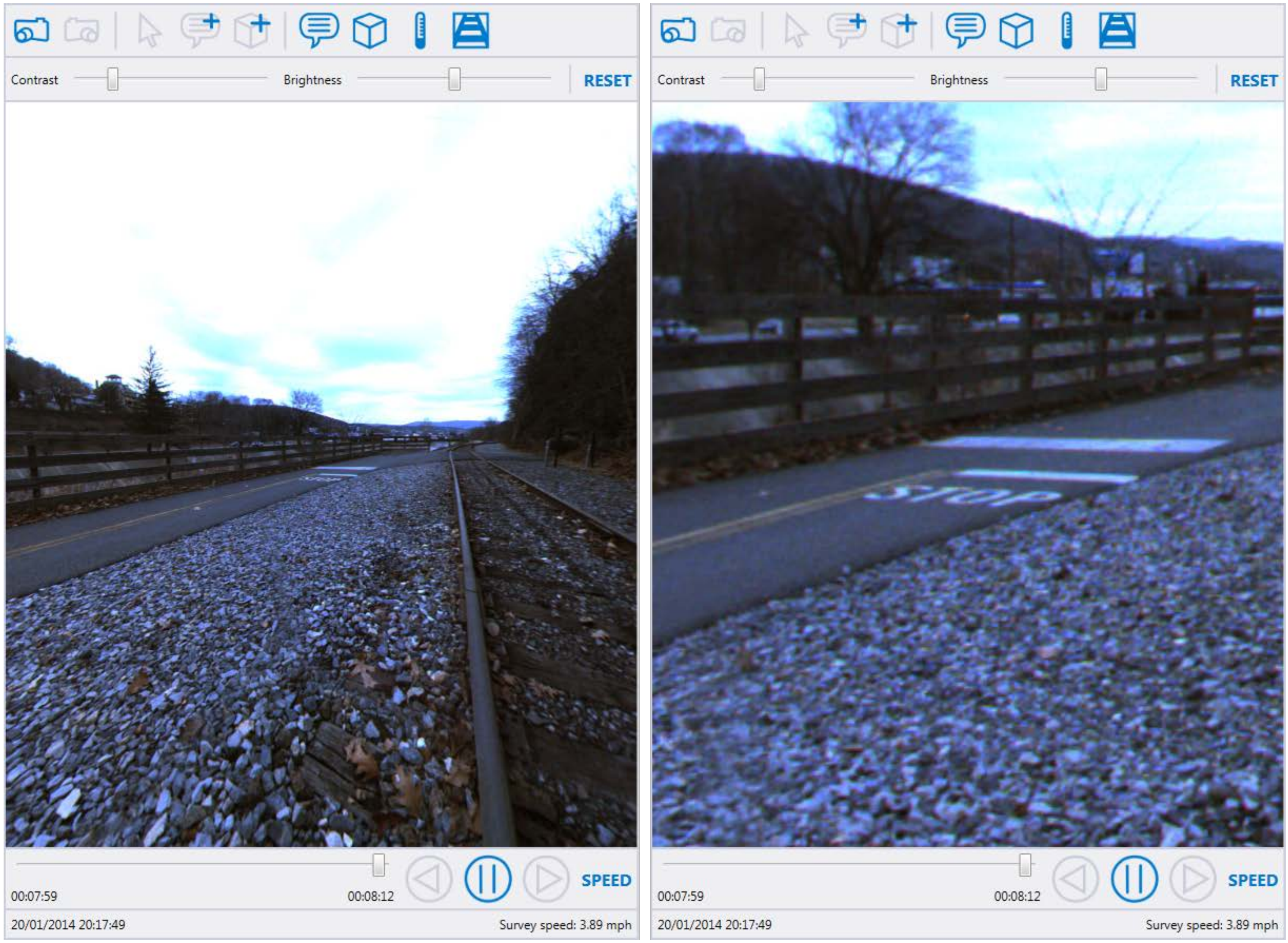


Figure 13. Additional WMSR context views.

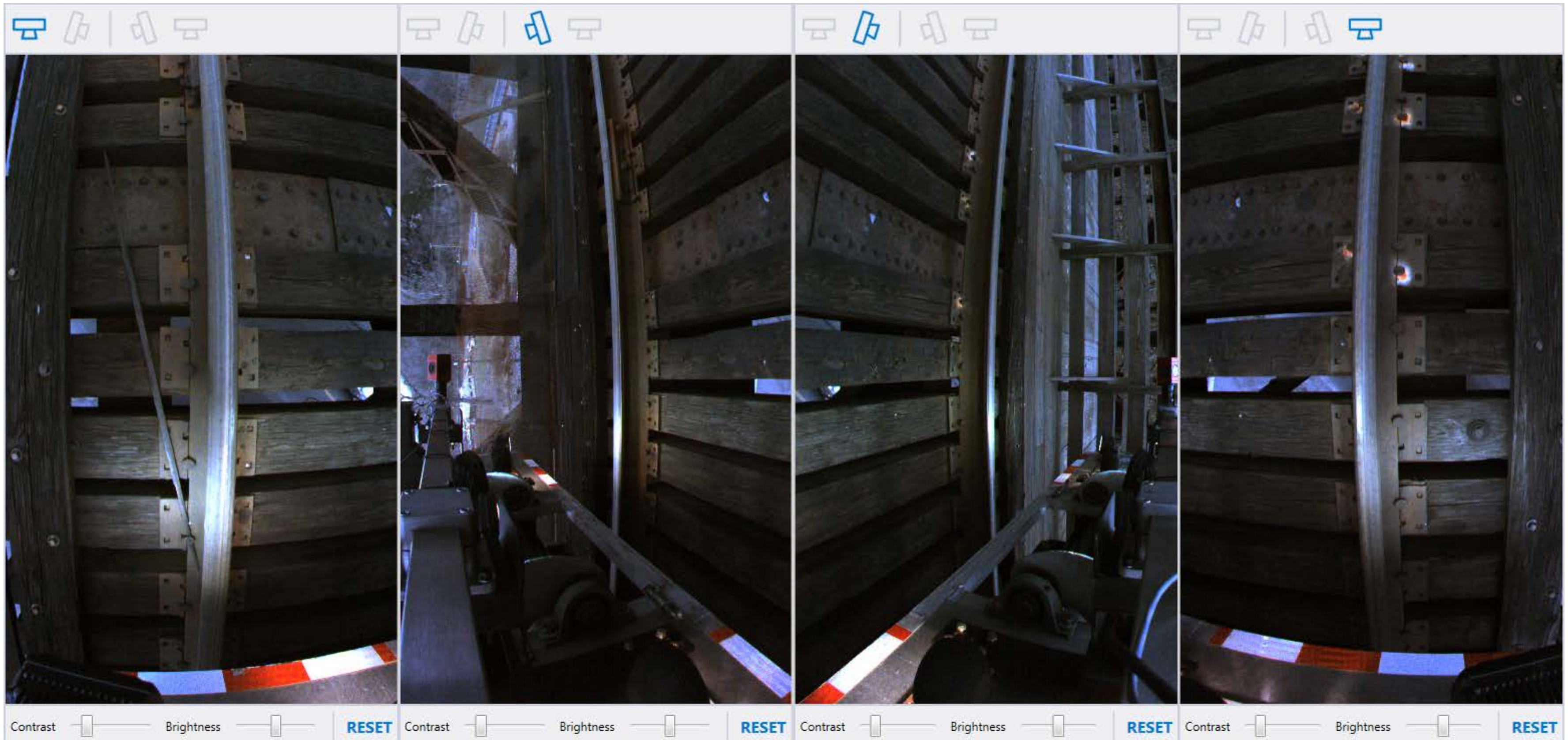


Figure 14. Track views while crossing a bridge.

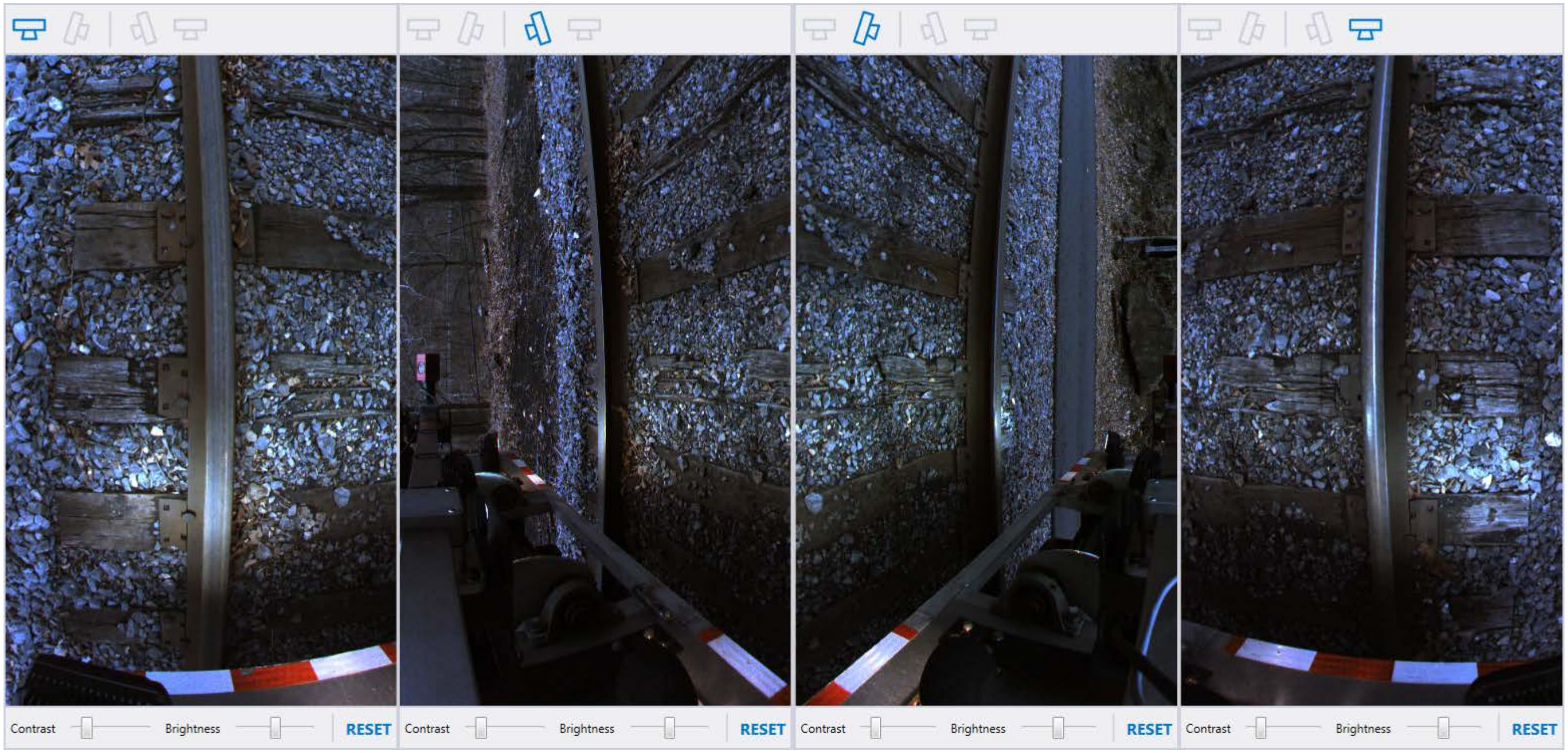


Figure 15. Track View over ballasted section.

Figure 24. Zoomed view of rails WMSR.

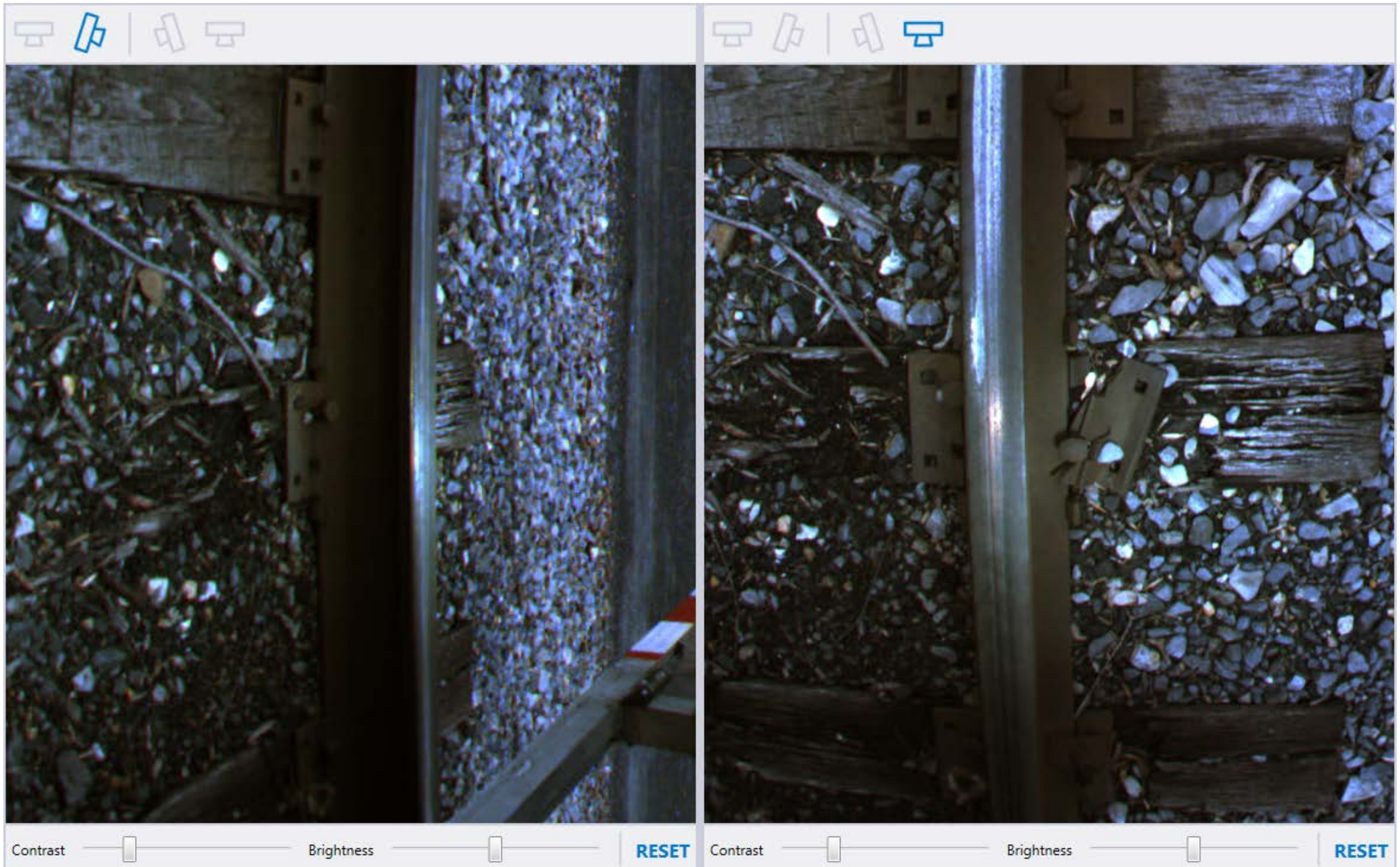


Figure 16. Zoomed view revealing broken tie plate.

3.1.3 Geo-location Results

The system accurately located the cameras for each frame in the survey videos. Figure 26 shows raw GPS readings and the output from the GPS-Egomotion fusion process from one survey run. Egomotion algorithms performed well with the WMSR data and successfully interpolated regions where GPS is lost.

Figure 17. Raw GPS readings (left), GPS-Egomotion fusion (right).

3.1.4 GPS Objects

Geo-located video enables the accurate creation and display of GIS objects. Figures 27 and 28 show the creation of survey note and tracked object GIS objects, respectively. The figures also compare the derived geo-locations against the locations that are observable on satellite imagery. Figure 29 shows what occurs when an object is tracked within the view. The user can save the imagery, the location, name, and a description of the object to the database for review at a later date.

GIS objects are observable on both the map interface and as overlays on video data. These video overlays are also interactive. Figure 30 shows a series of frames from a survey video with a number of GIS objects visible in the frame. These objects remain fixed in space as the camera approaches. Figure 31 depicts the user selecting a GIS object in the distance and asking the video to navigate to the nearest frame. Figure 32 shows the result of this action; the video skips ahead and adjusts the field of view to focus on the object of interest.

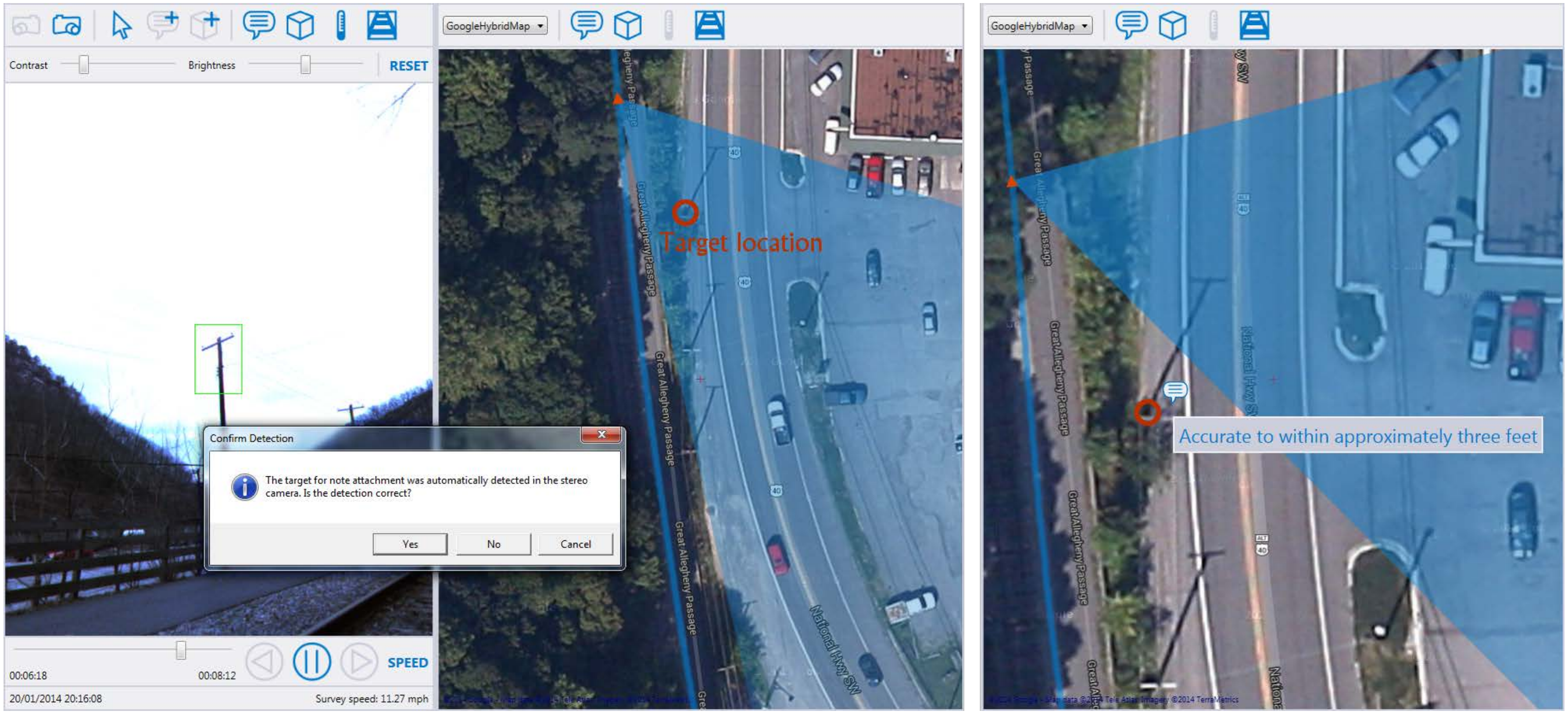


Figure 18. Survey note creation and location accuracy.

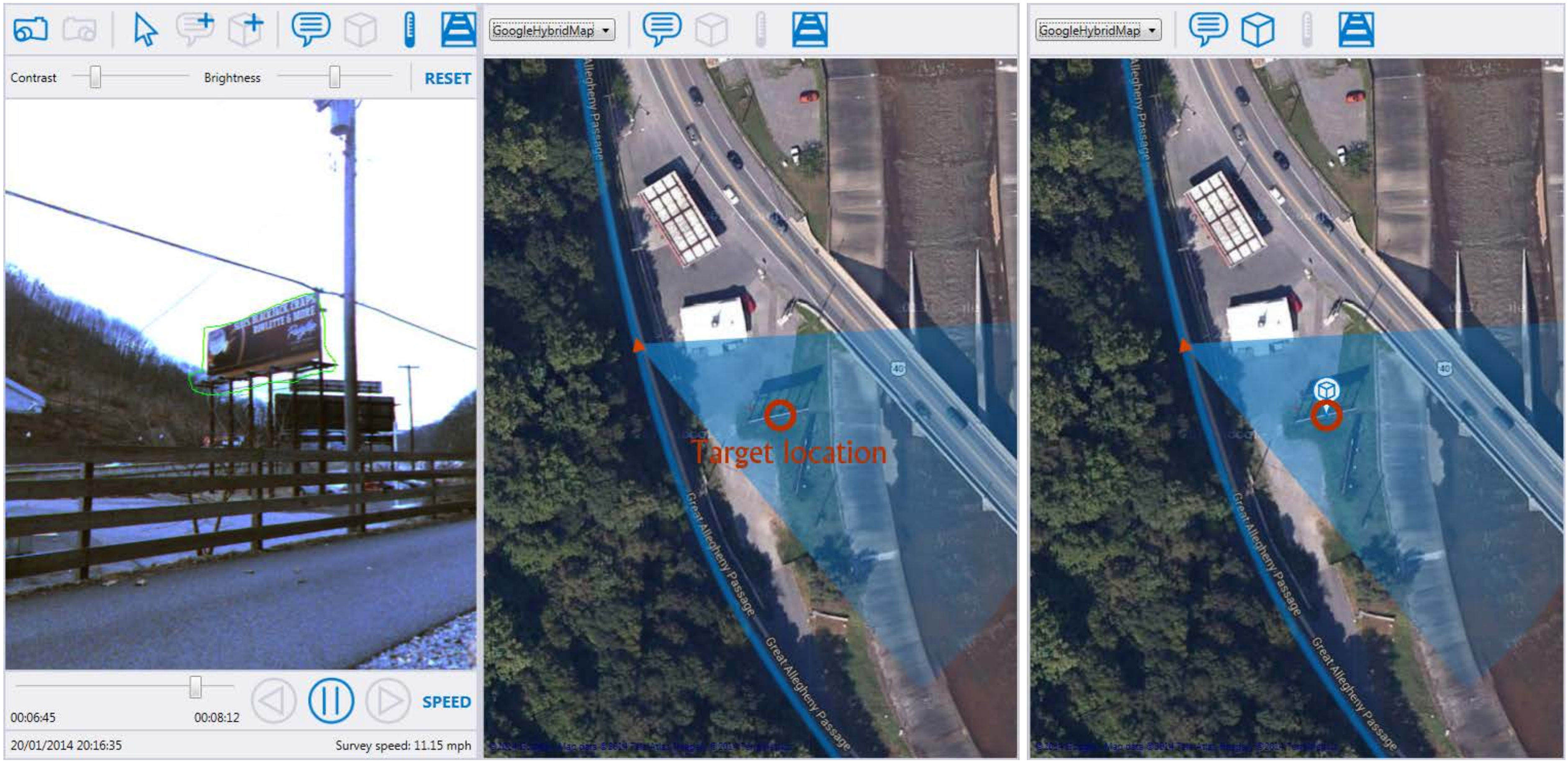


Figure 19. Object tracking.

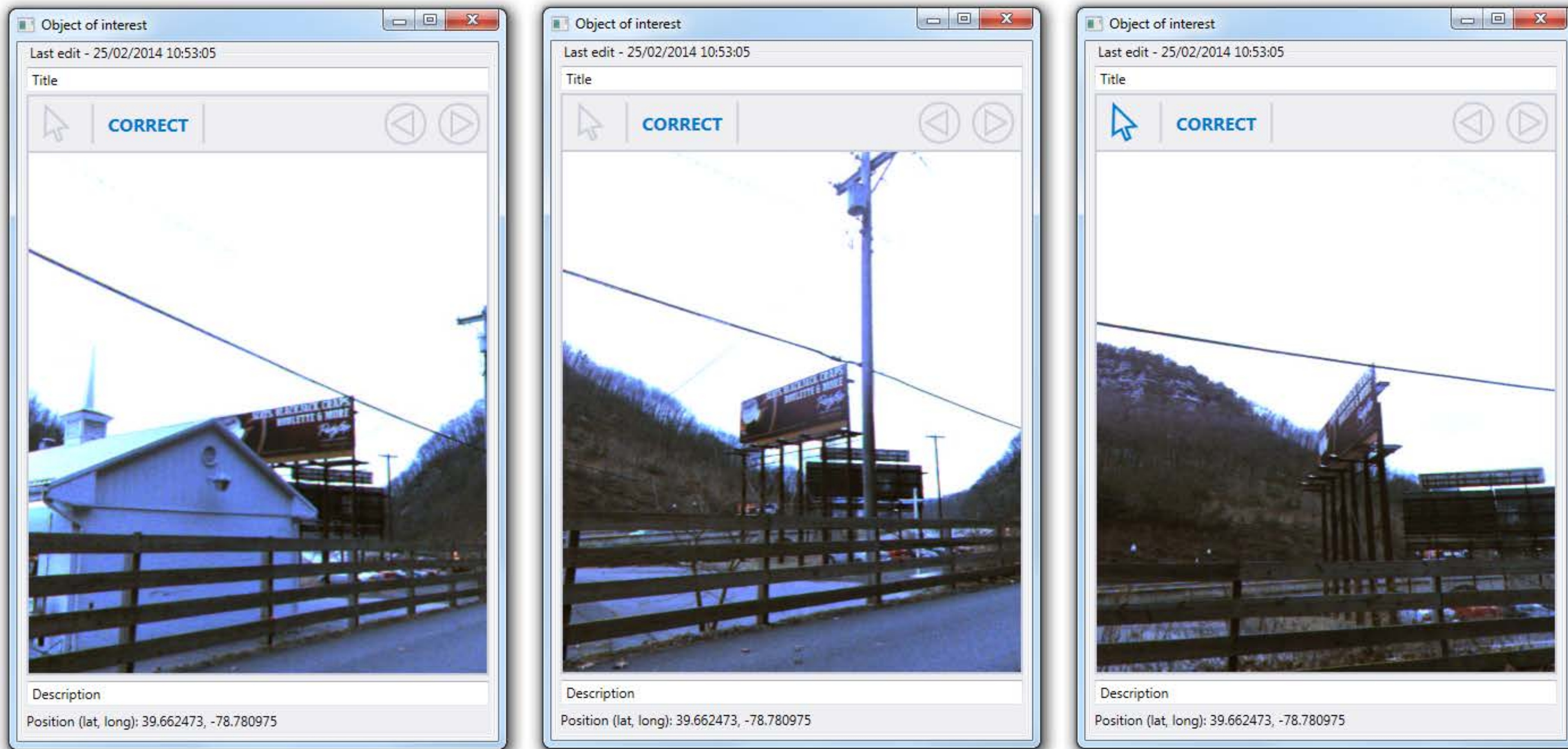


Figure 29. Object tracking images.

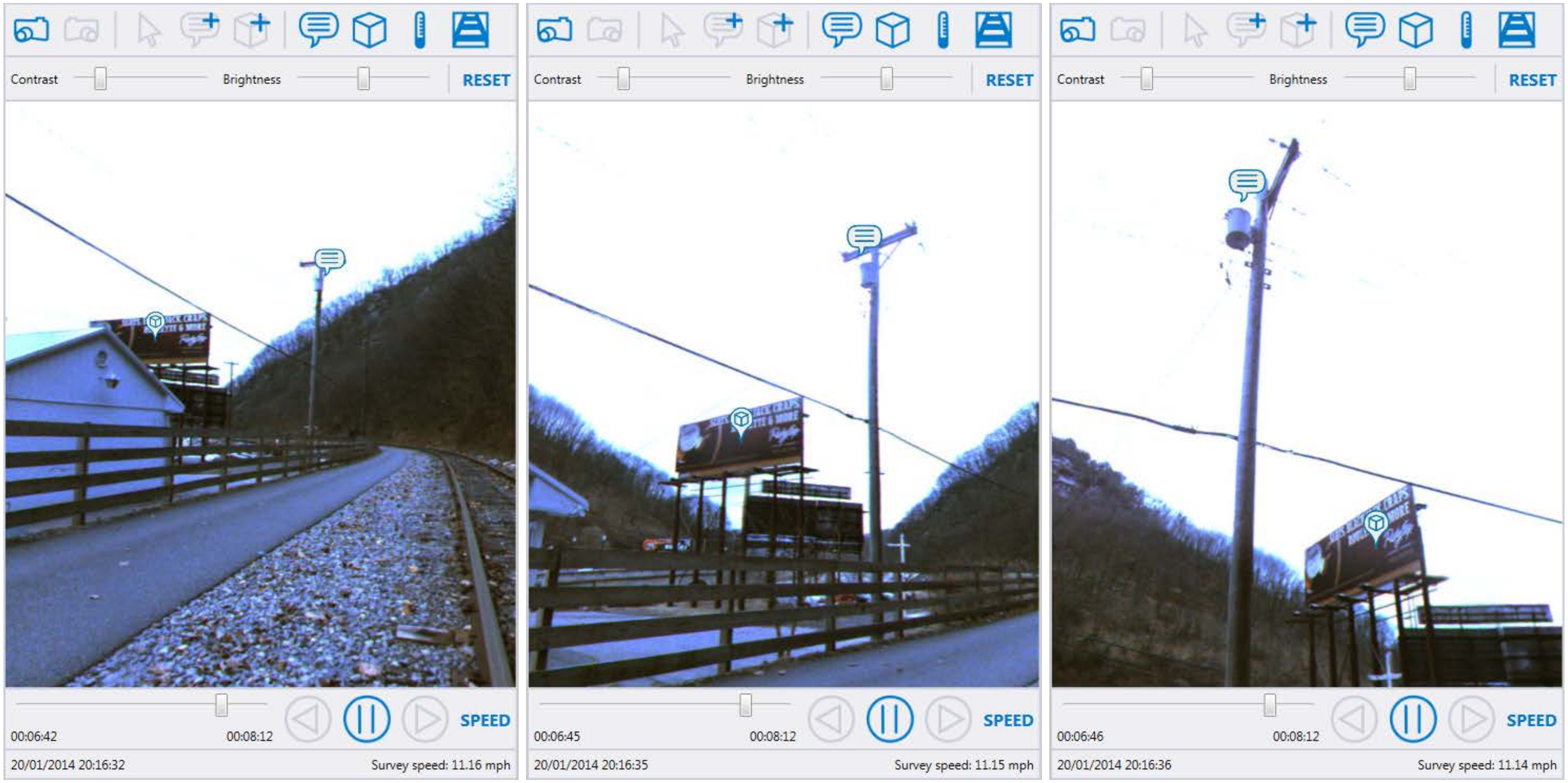


Figure 20. Frames showing fixed GIS overlays.

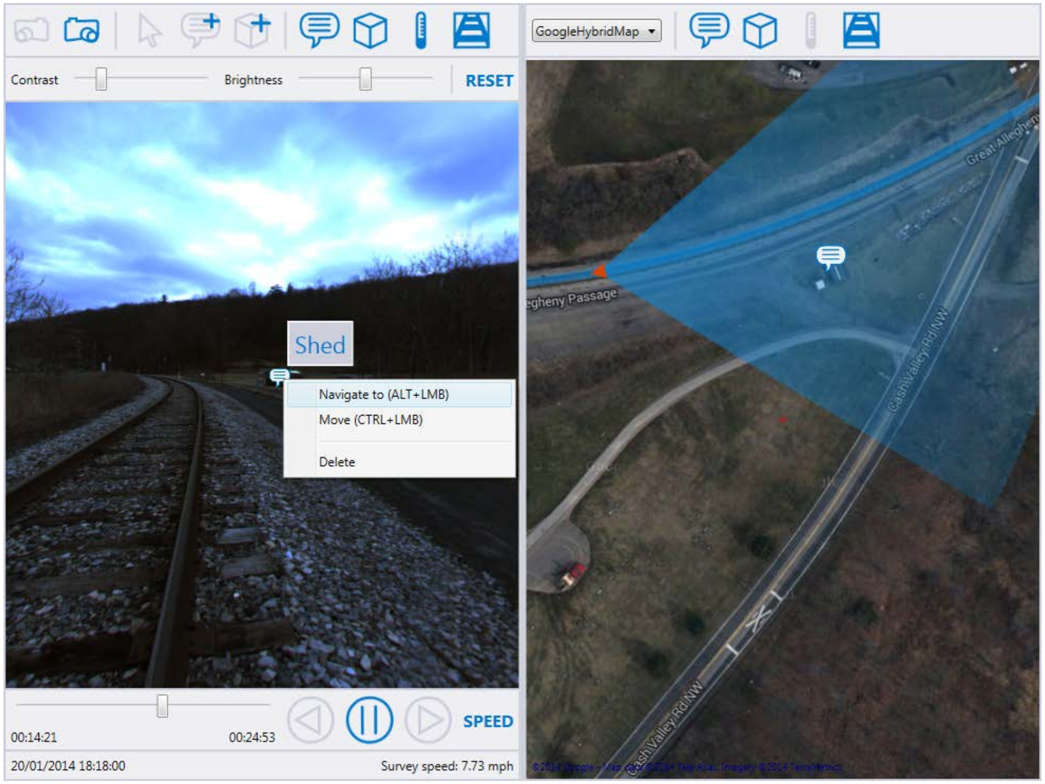


Figure 21. GIS object selection and navigation.

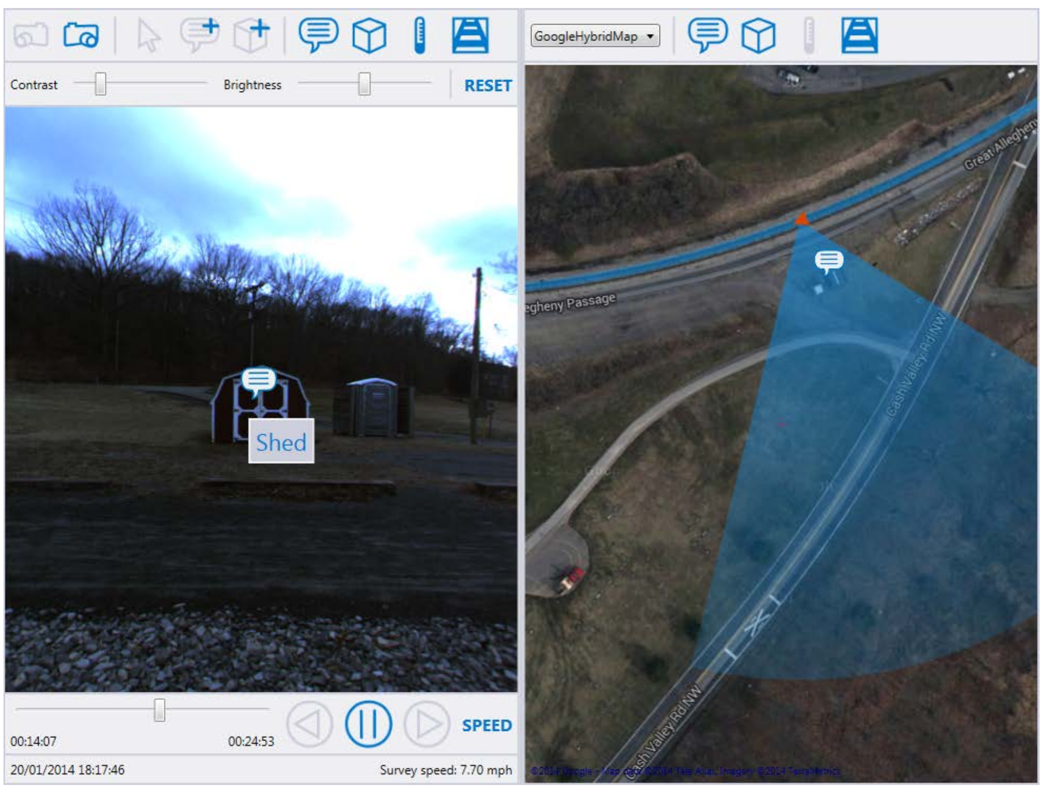


Figure 22. Video advance to selected GIS object.

3.1.5 *Incorporating External Data*

The TTVS GUI provides the user with a natural and intuitive way to navigate large quantities of data. This virtual environment is powerful because it not only displays data generated internally but also displays data imported from external sources. The TTVS virtual environment binds together geo-located survey data from a number of survey processes and presents it in a consistent and easy to understand manner. Gathering multiple survey data into one environment provides greater insight to the user than if the data is in isolation and enables him or her to quickly gain a comprehensive understanding of track condition.

In order to demonstrate this functionality, Createc was provided with a series of geo-located rail temperature measurements and joint bar images generated during other maintenance activities.

3.1.5.1 *Rail Temperature Measurements*

The rail temperature measurements were loaded into the GIS-enabled database so they could be used within the GUI. For each temperature measurement, the GUI displays a marker that is color coded to match a temperature gauge displayed as an overlay (see Figure 33). Hovering over a marker reveals a precise temperature and the date that the measurement was taken. Figure 34 shows these markers as video overlays.

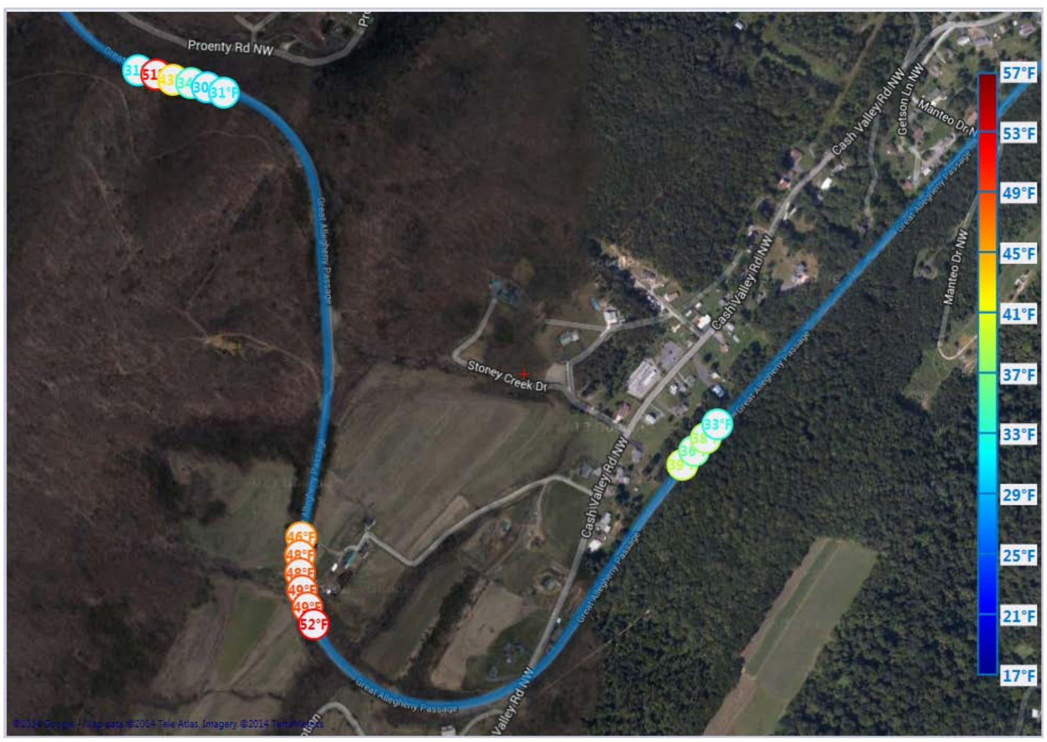


Figure 23. Rail temperature measurements imported from an external data set.

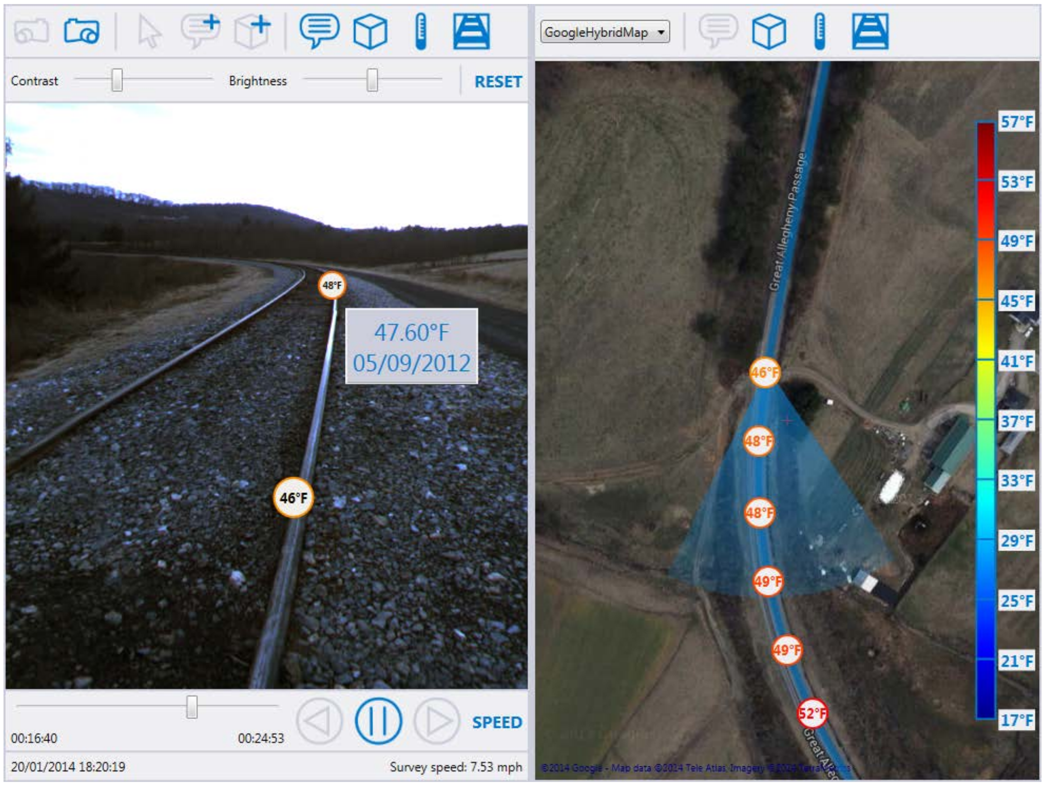


Figure 24. Temperature markers visible as video overlays.

3.1.5.2 Joint Bar Images

Details on each joint bar image were loaded into the GIS-enabled database for use within the GUI. The GUI displays a marker on the map and a video overlay for each image within range (see Figure 35). Selecting a marker opens a pop-up window containing the image along with its title and description (see Figure 36).

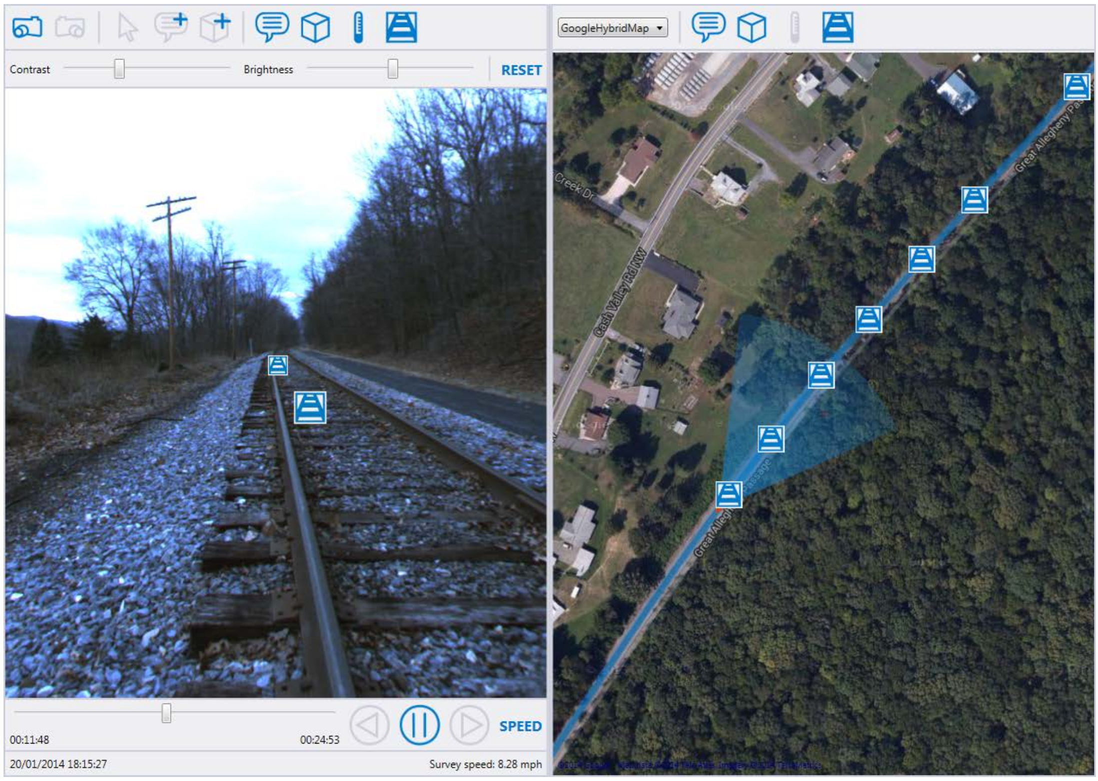


Figure 25. Joint bar image with GIS markers.

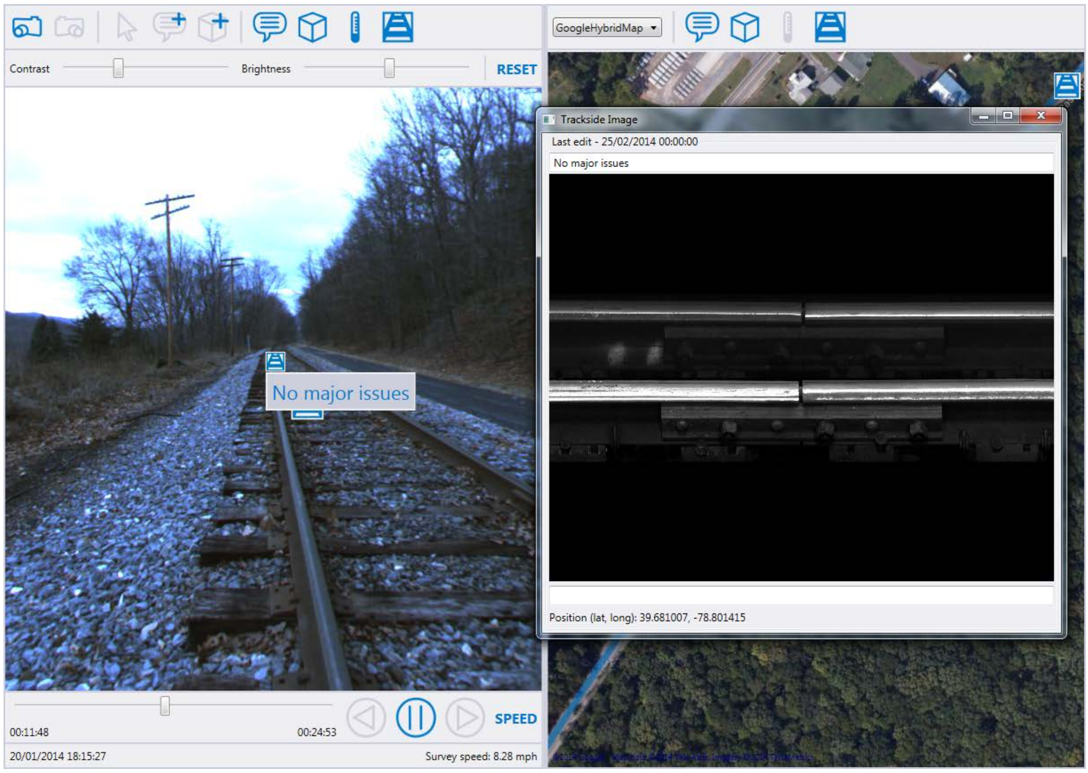


Figure 36. Joint bar image from external data source.

3.2 Bay Line Railroad

A subsequent demonstration test was completed on the Bay Line Railroad (BAYL) near Panama City, Florida. This test was conducted in tandem with a geometry survey. This was as an opportunity to demonstrate the system's ability to integrate with existing survey systems and assess performance under different survey conditions.

3.2.1 Survey Execution

ENSCO undertook the Panama City trial between February 18th and February 22nd 2014 and gathered data from four runs over a 30 mile track segment. Weather conditions were favorable compared to the WMSR survey, enabling survey speeds to reach 28 mph (150 percent design speed) without any observable motion blur. The camera system was mounted slightly lower for the Panama City trial at 0.8 m above track level.

The TTVS anti-vibration system was designed for a maximum survey speed of 20mph and it was found to be insufficient for the higher survey speeds reached during this trial. This did not affect the quality of survey data that was gathered. Createc will redesign this anti-vibration system in future TTVS systems.

3.2.2 Example Imagery

Due to favorable light conditions, the imagery obtained from the Panama City trial was of a higher quality than the WMSR dataset. Images are bright and crisp despite survey speeds reaching 28mph.

Figures 37 and 38 demonstrate the system's ability to zoom into regions of images via the context view of the GUI. Figures 39 and 40 depict images from the four selectable views in the track view window of the GUI. Figures 41 and 42 show how the system can zoom down into track images and focus on an area of interest.

Figure 43 shows three of the views that resulted from tracking a level crossing signal. The associated GIS object is observable on both the map interface and as an overlay on video data; see Figure 44. Figure 45 shows a series of frames from a survey video with a number of GIS objects visible in the frame. These objects remain fixed in space as the camera approaches them.

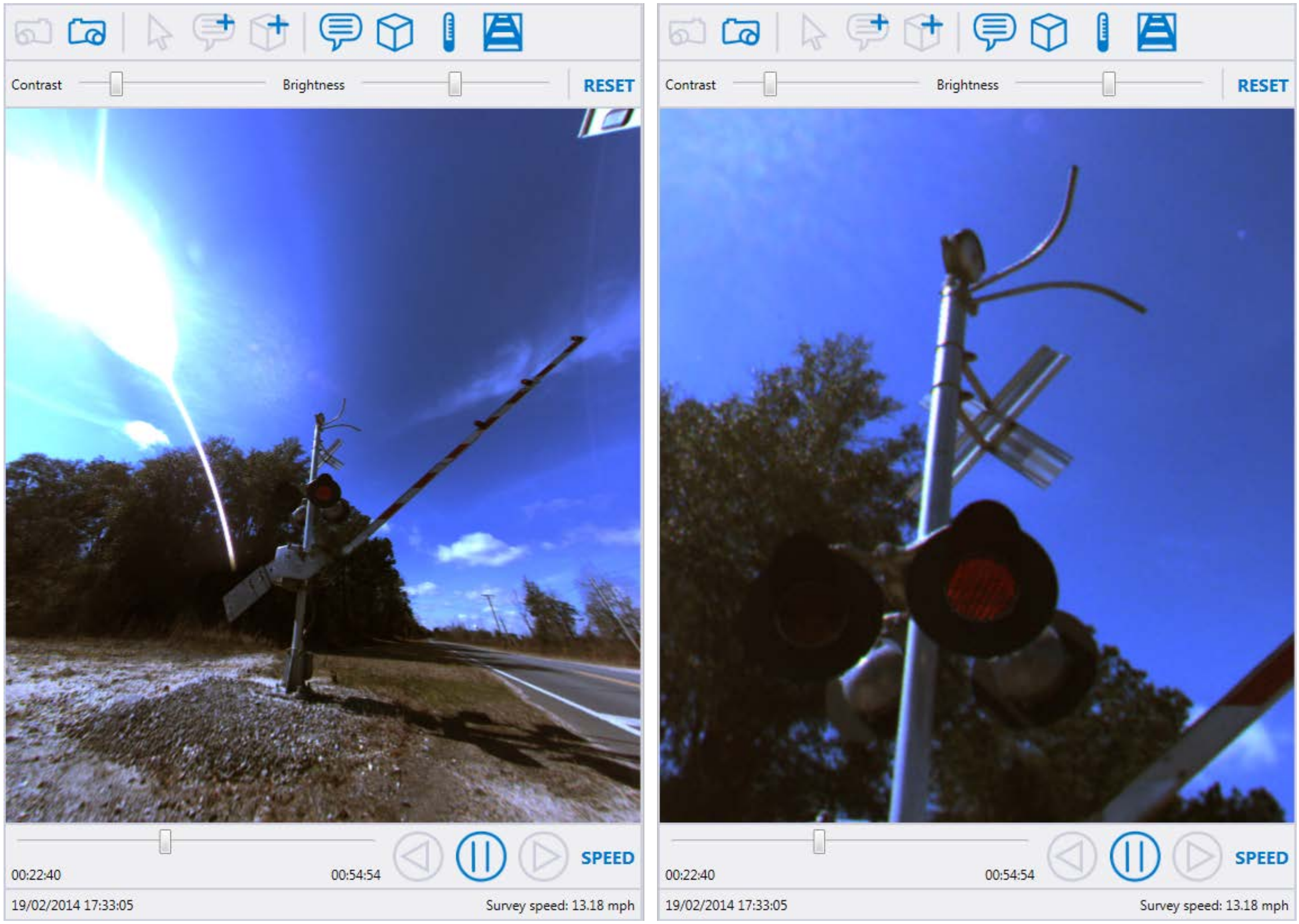


Figure 26. Context view screenshots of Bay Line data.

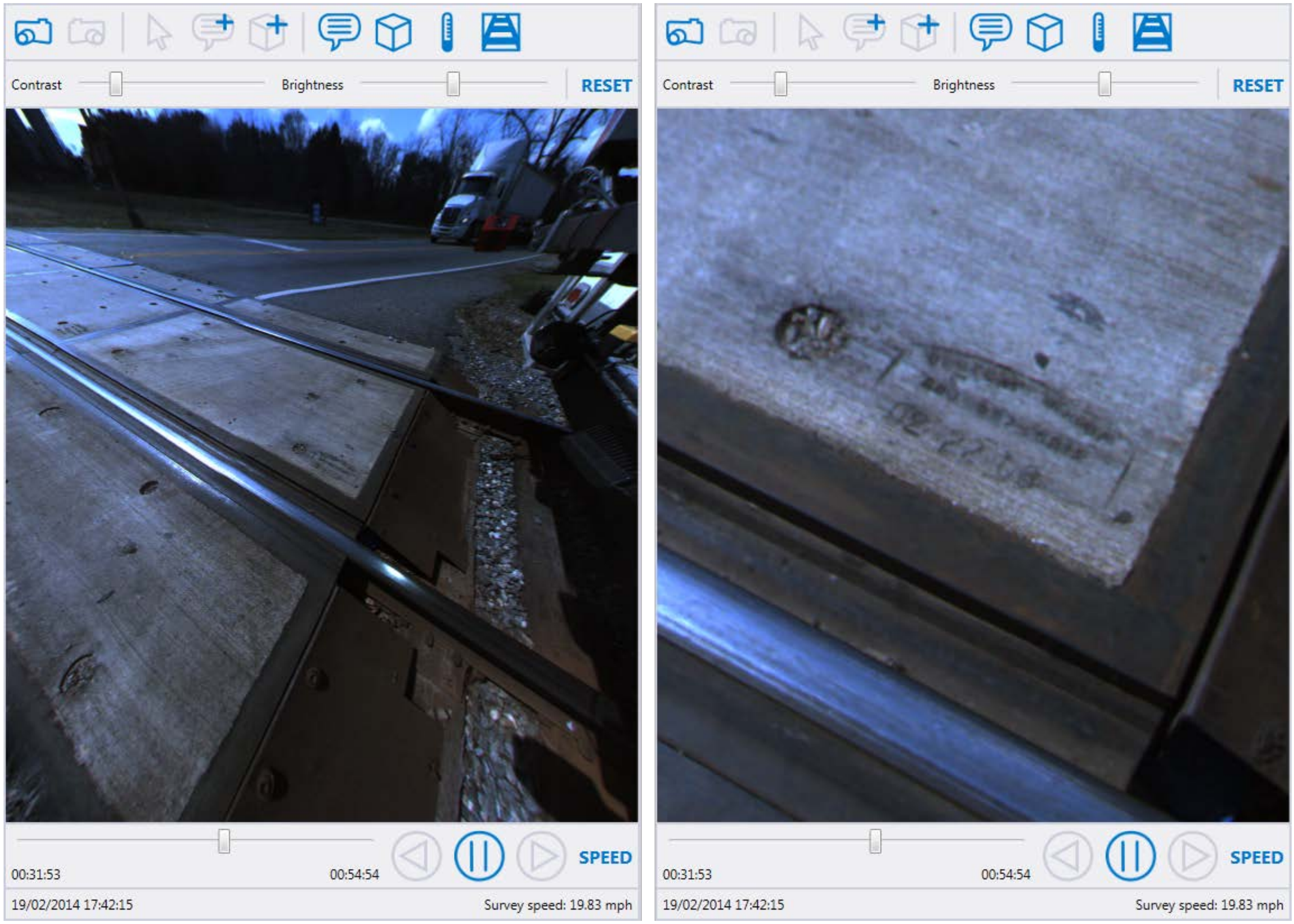


Figure 27. Additional context view screenshots.

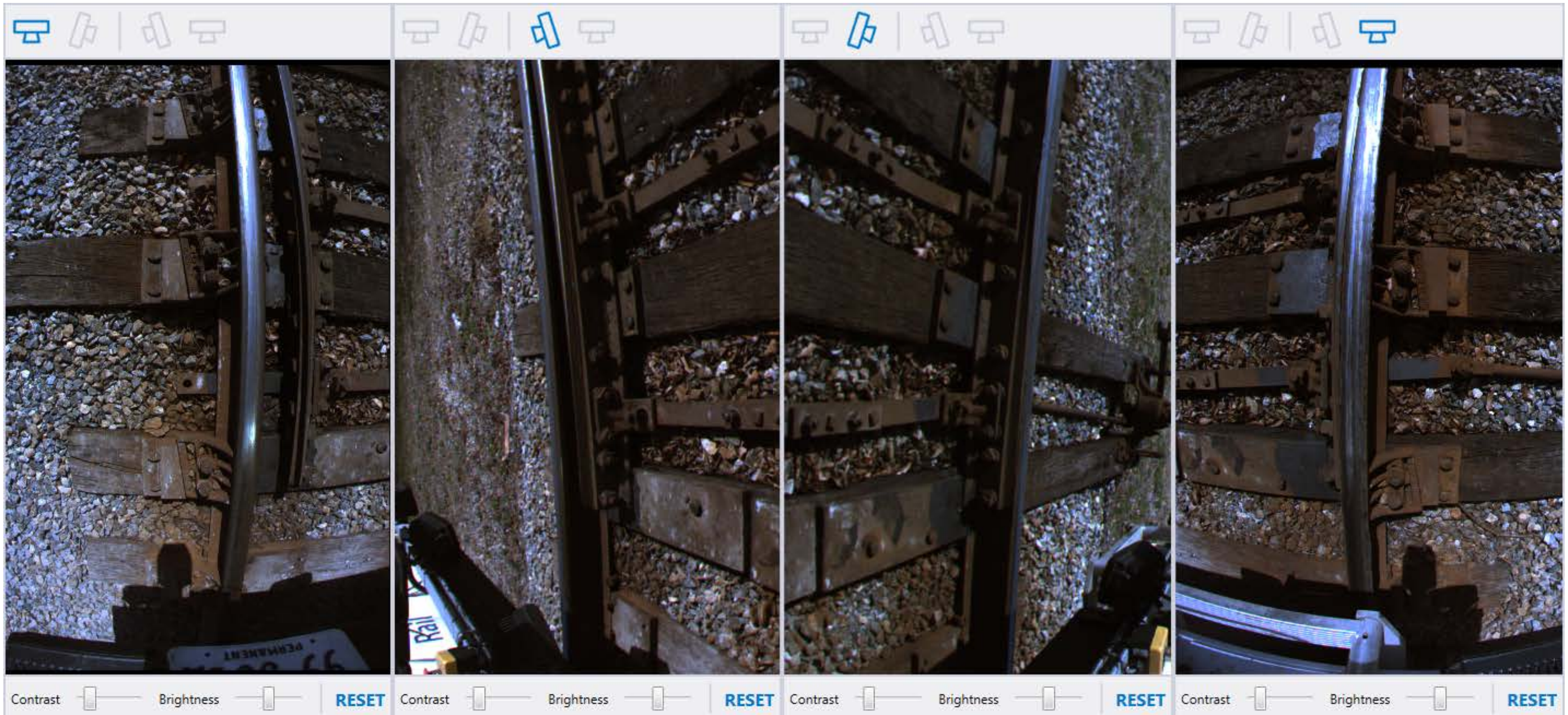


Figure 28. The four track view images.

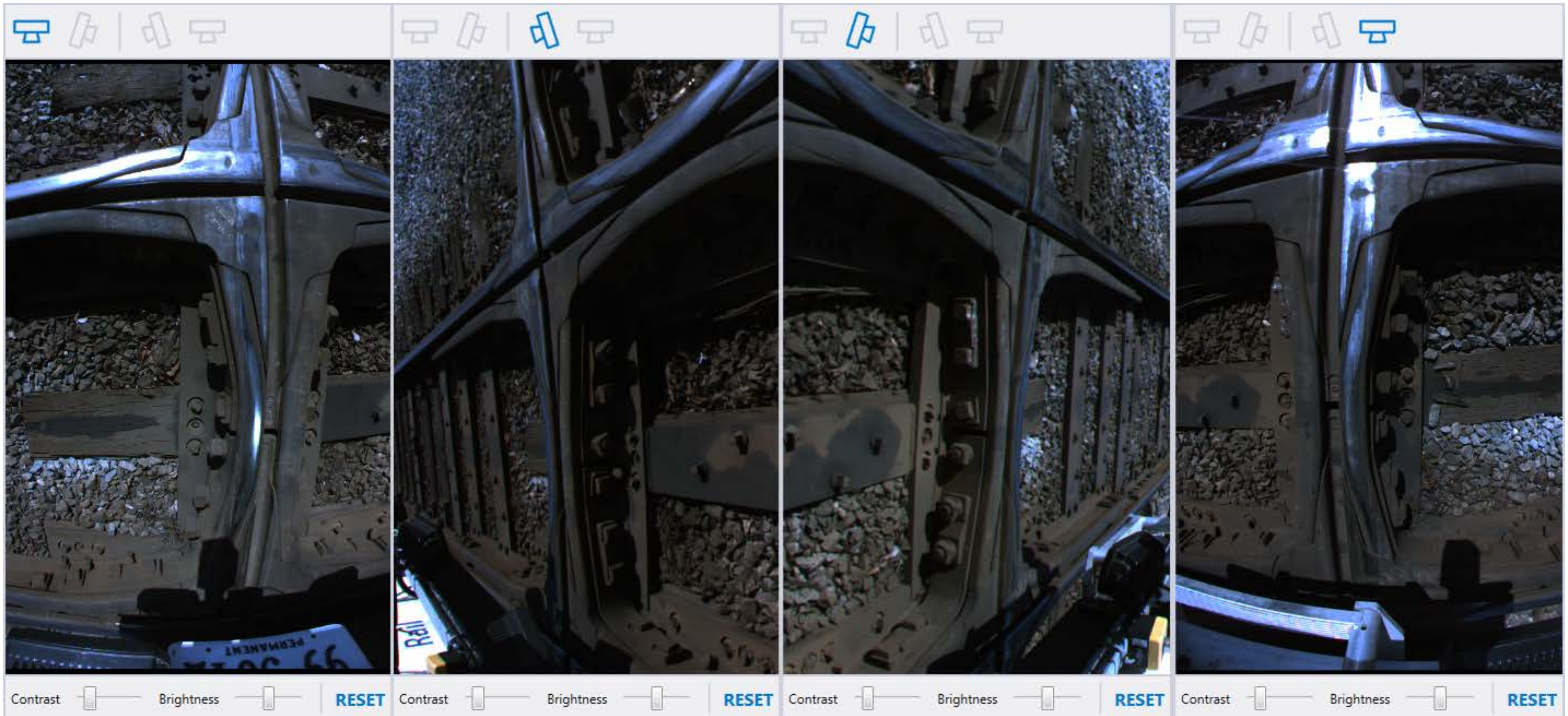


Figure 29. Additional track view images.

Figure 30. Zoom view showing failed crosstie.

Figure 31. Zoom view revealing rail wear.

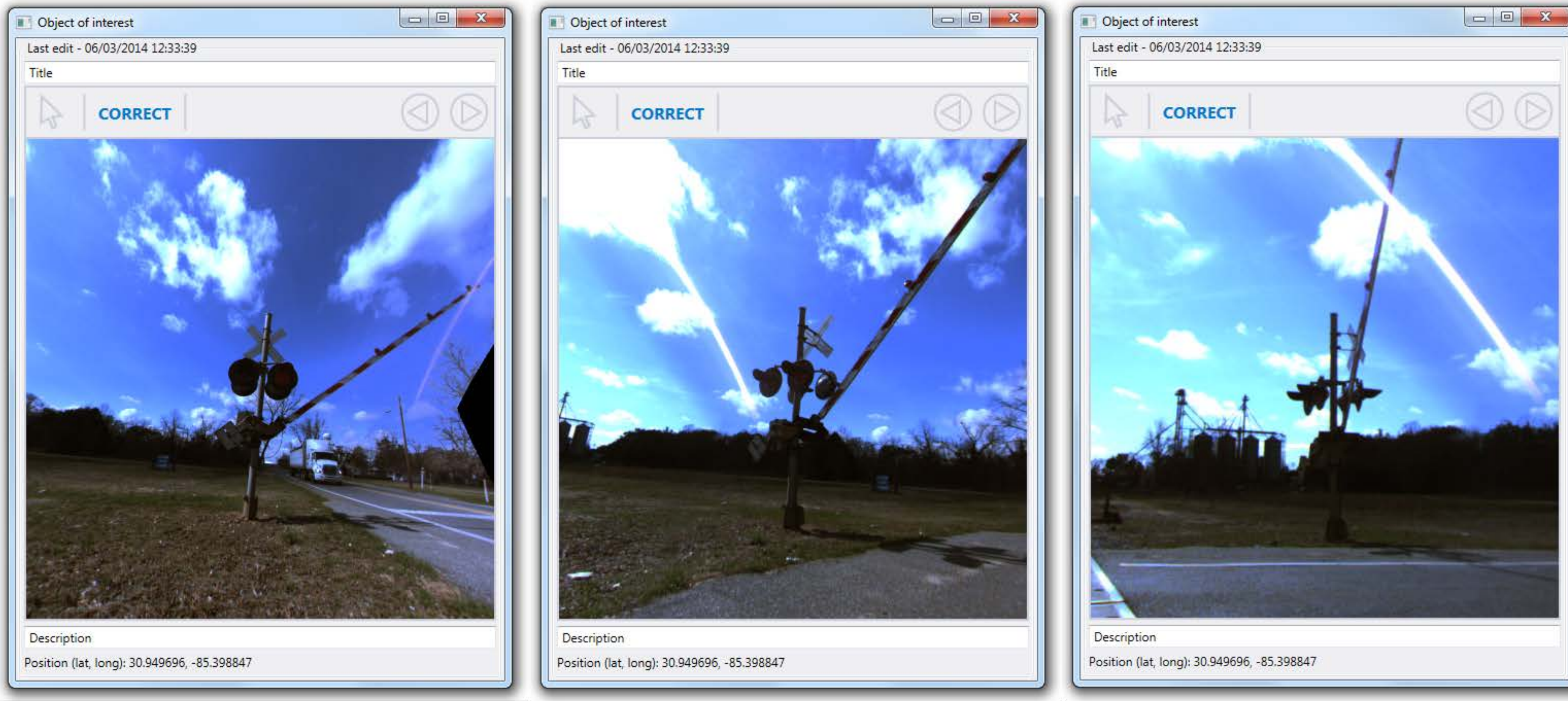


Figure 32. Object tracking images.

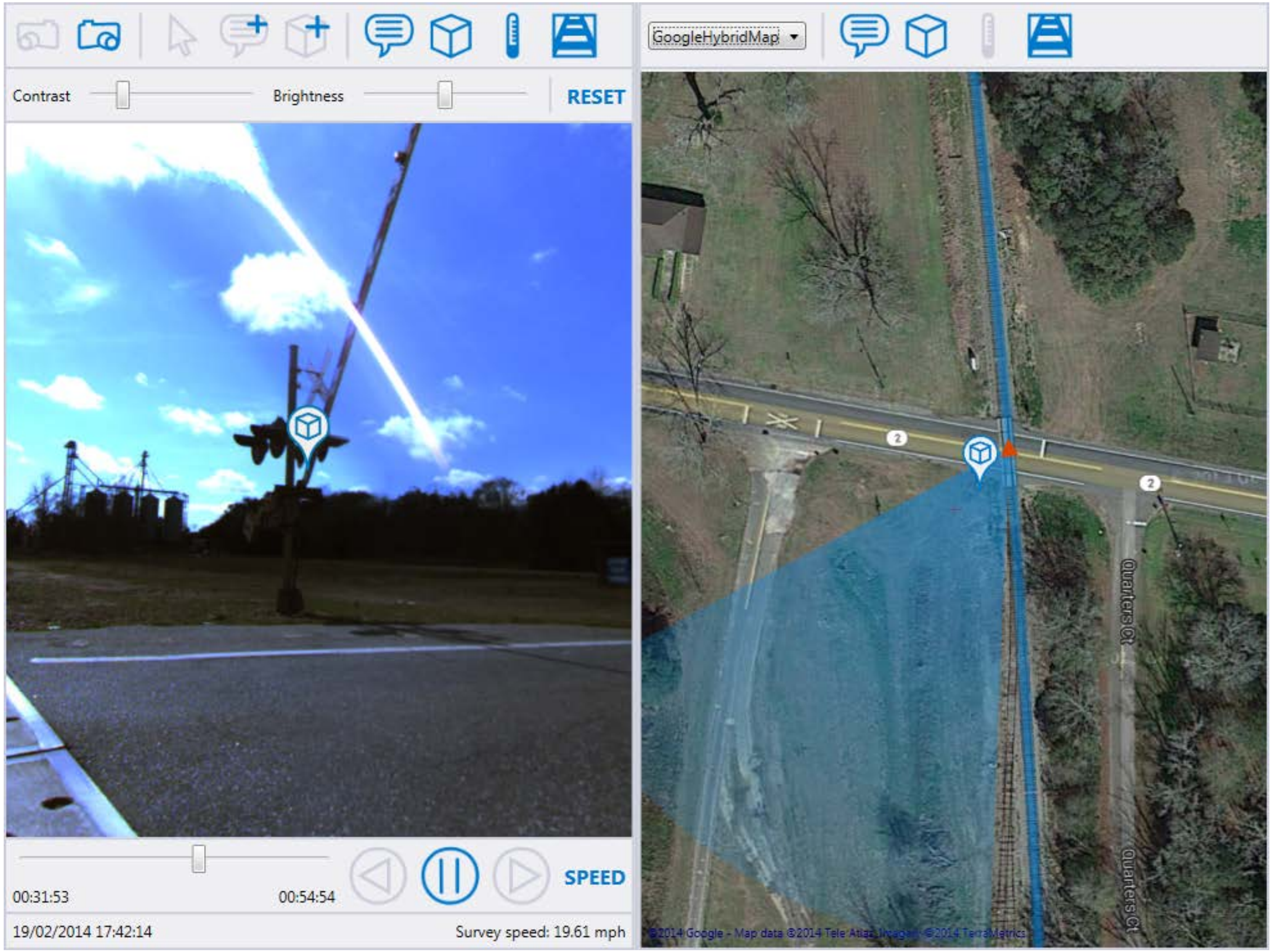


Figure 33. The GIS object as an overlay on context and map views.

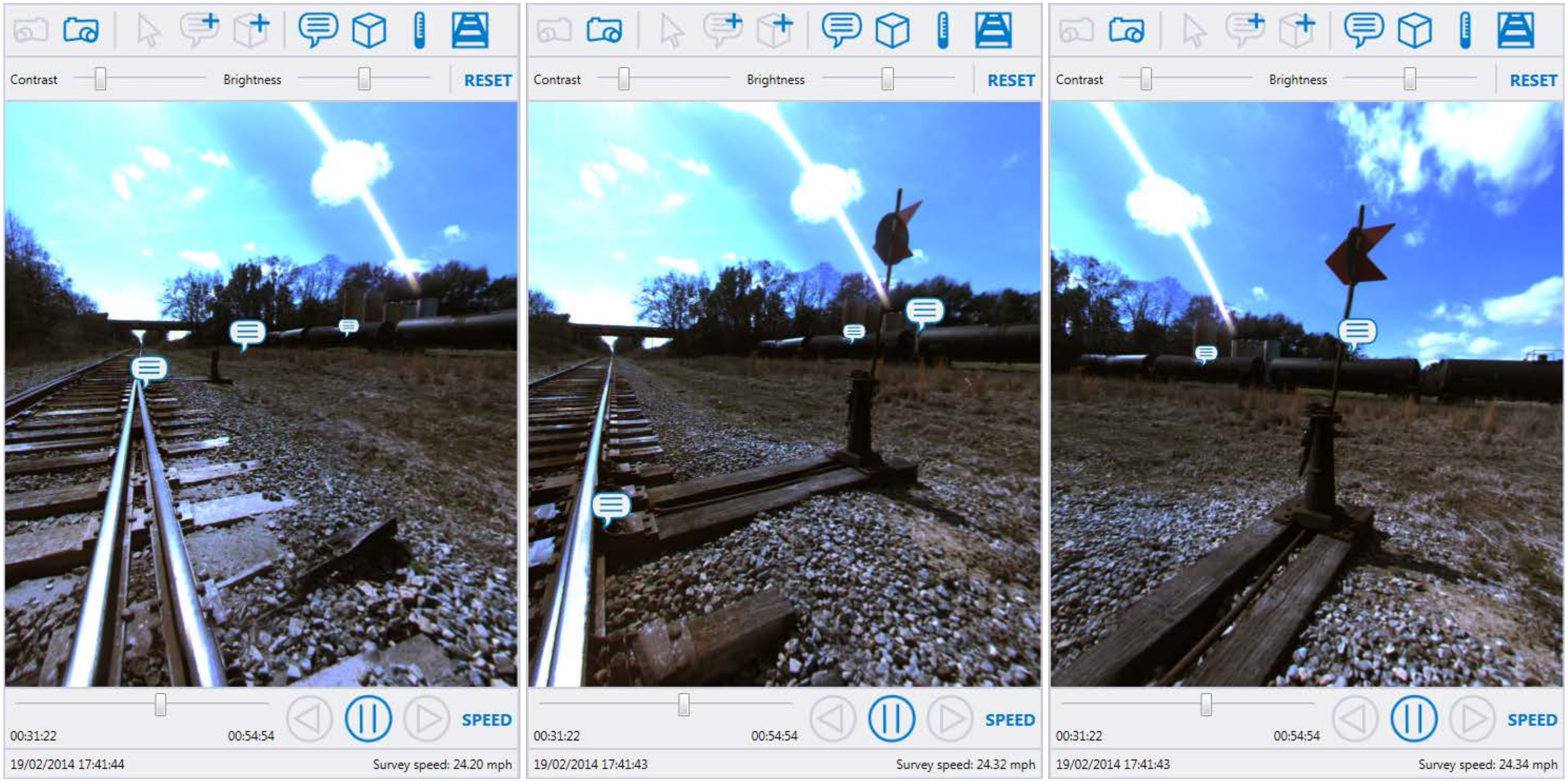


Figure 34. GIS objects tracking through video images.

4. System Performance Evaluation

This section of the report assesses the performance of the system and identifies areas for further development.

4.1 Image Quality

Both TTVS trials demonstrated the ability of the system to produce track imagery at a resolution of 1 mm or better. The best track imagery was attained during the Panama City trial as the camera system was mounted slightly closer to the track, increasing resolution (see Figure 46).

As anticipated, the general quality of images captured by the system is weather-dependent. Significant cloud cover reduces light, increasing image noise. Figure 47 contains images typical of the range in quality attained during the trials. The lower quality image was taken under heavy cloud cover, while the higher quality image was taken on a bright day.

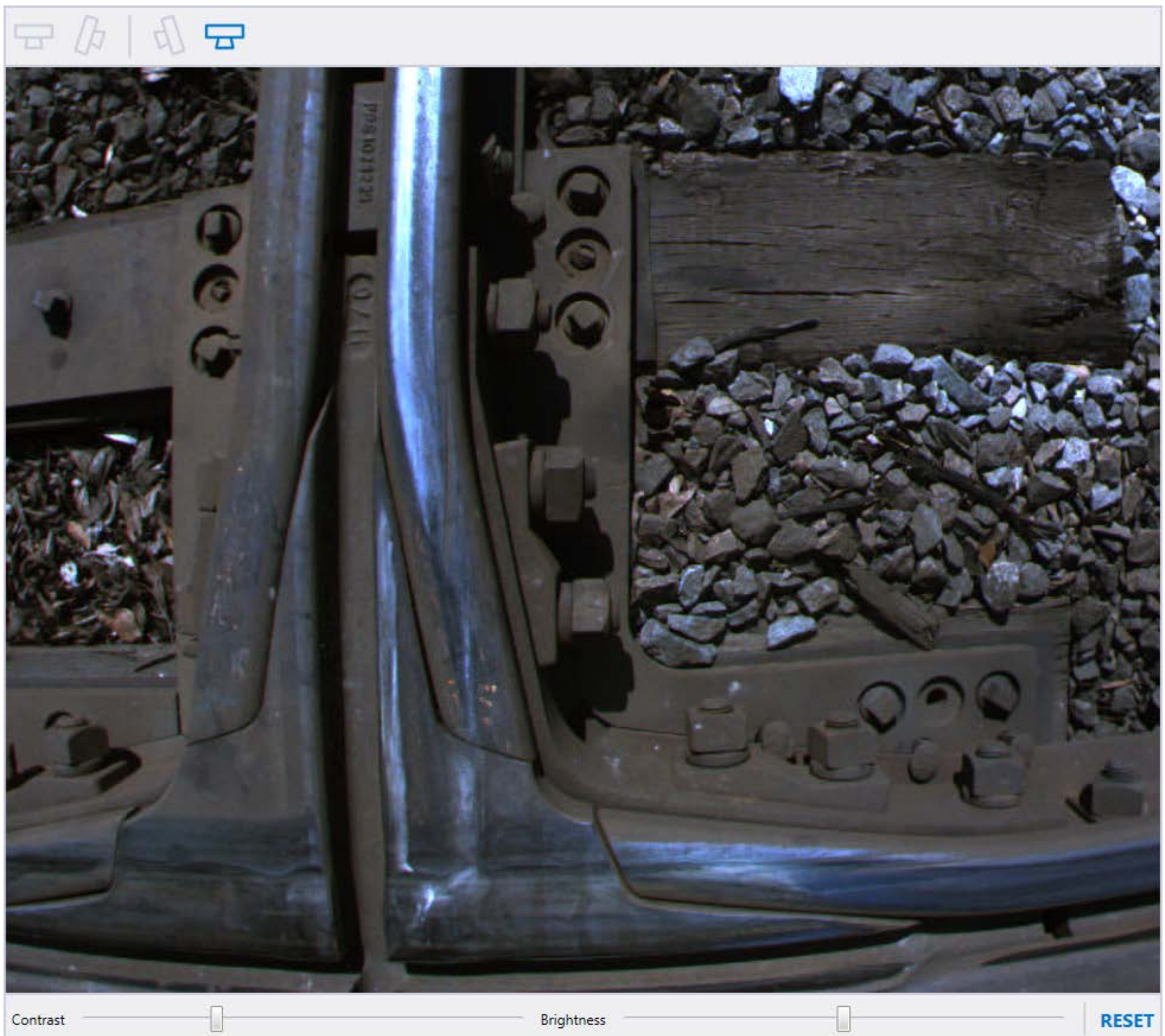


Figure 35. High quality track imagery.

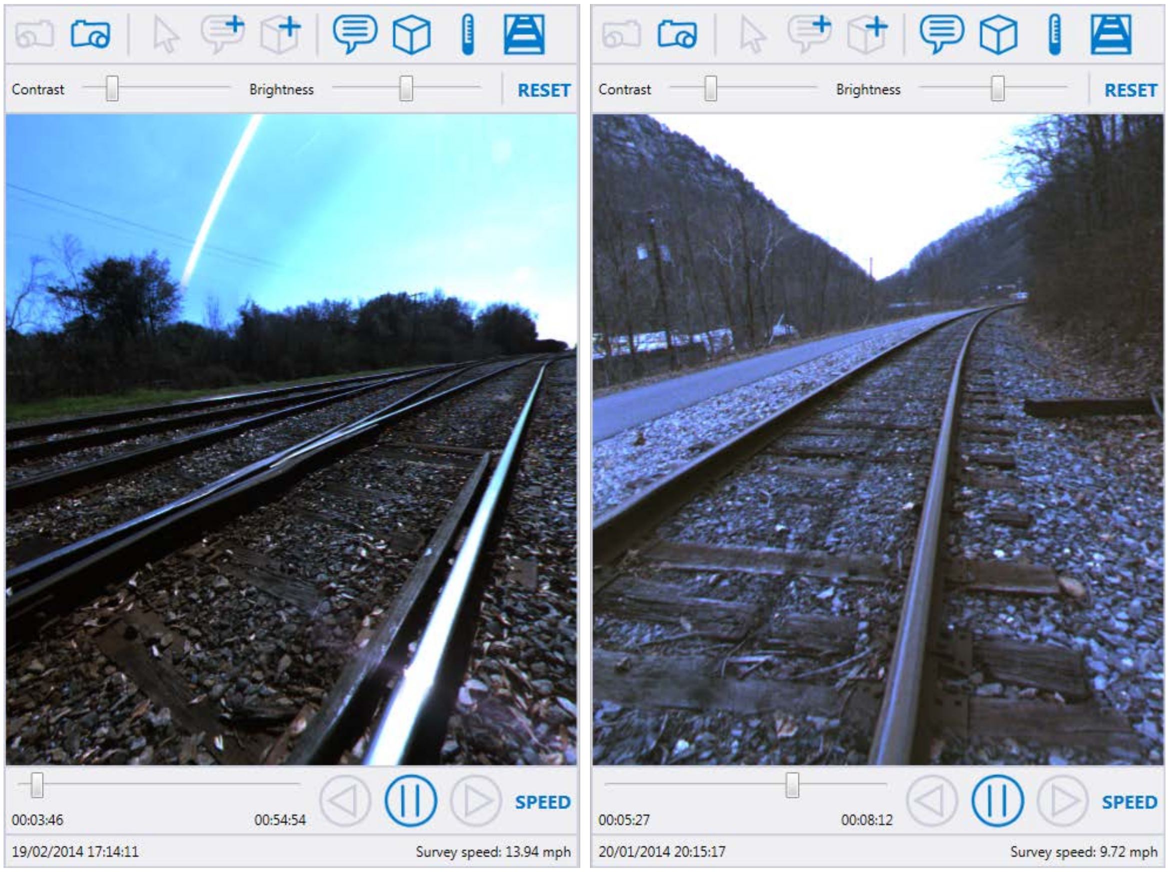


Figure 36. Weather-affected image quality.

Adverse weather conditions introduced the issue of motion blur that was caused by long exposure times. Researchers overcame this problem by overriding the camera's automatic exposure settings and manually specifying a shorter exposure time. This fix comes at the expense of increased image noise. However, this noise does not affect the functionality of the system. Future TTVS systems will be able to automatically calculate optimum exposure settings for various lighting conditions.

4.2 Geo-Location

Combining Egomotion with GPS accurately geo-locates survey footage. Typically, the errors observable during the trials were less than 1.5m. This level of precision is more than enough for navigating video footage; the system can accurately move to a location selected on the map or to a GIS object created by another user.

The creation of GIS objects through the interface is equally accurate. By examining satellite imagery, it is evident that the automatic geo-location of objects selected in the video is typically accurate to within 2 m; accuracy would improve further with better geo-location of video footage.

Geo-locating an object within the video only requires a single user interaction. Under some circumstances the system has problems obtaining a stereo lock on the target. This situation can occur when a target is very close to the cameras and or between the stereo pair, making the target appear significantly different to each camera. In these situations, the GUI prompts the user to select the object a second time.

Video footage is geo-referenced using a combination of GPS and Egomotion. Egomotion is a machine-vision technique for measuring camera motion directly from image data and it is capable of accurately measuring small movements and rotations. Egomotion and GPS together emulate an INS, providing an accurately geo-located sensor package that can be used as a platform for the addition of further sensors such as LIDAR or thermal imagery at a later date.

Geo-referenced footage allows the TTVS system to combine images with a geographic information system (GIS). A simulated aerial view can be generated enabling images to be associated with and navigated using a map view. GIS objects can be visualized as overlays within the virtual survey environment and they can be created using simple video mark-up tools.

In addition to viewing data from the train's eye view, object tracking algorithms enable the user to switch focus onto a track-side object and navigate around it, viewing it from different directions (as shown in Figure 48). The tracked object is exported as a GIS object and saved to a database for review at a later date, which provides a powerful tool for change detection to users.

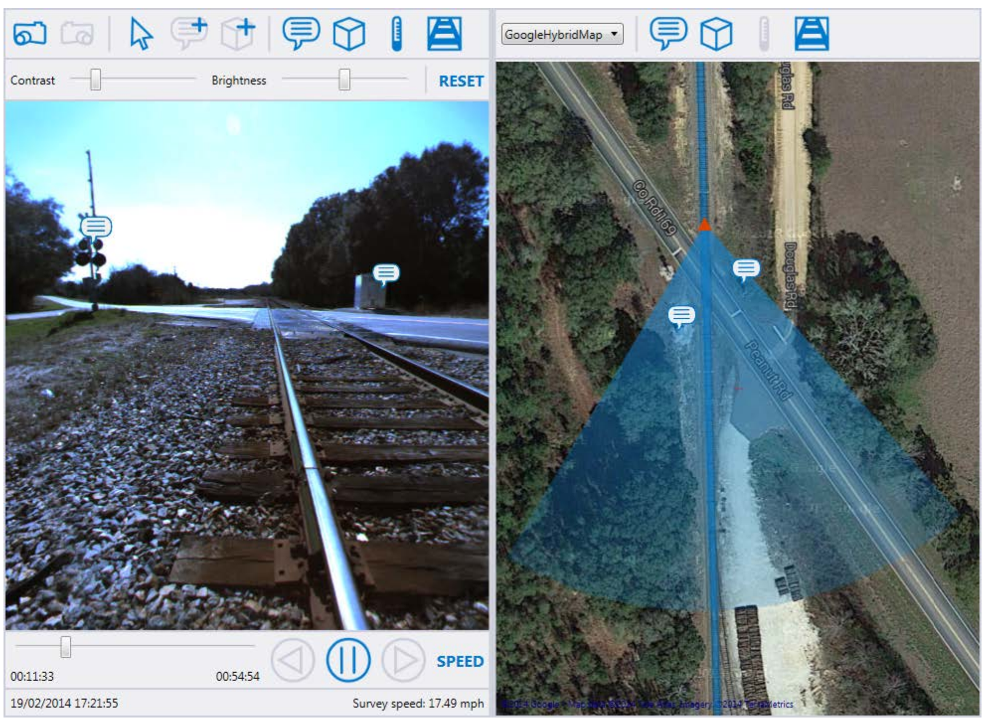


Figure 48. Navigating geo-referenced locations in the virtual environment.

4.3 Object-Centric Video Navigation

The system's ability to select an object from within the video and automatically extract all frames in which it is visible is a powerful track and wayside inspection tool. The object tracking algorithms that enable this functionality perform well and they successfully track objects in the vast majority of cases. For the rare occasions when tracking fails, the user can employ a manual correction tool. The user can save and recall tracked targets to and from a GIS database, providing a powerful tool for change detection.

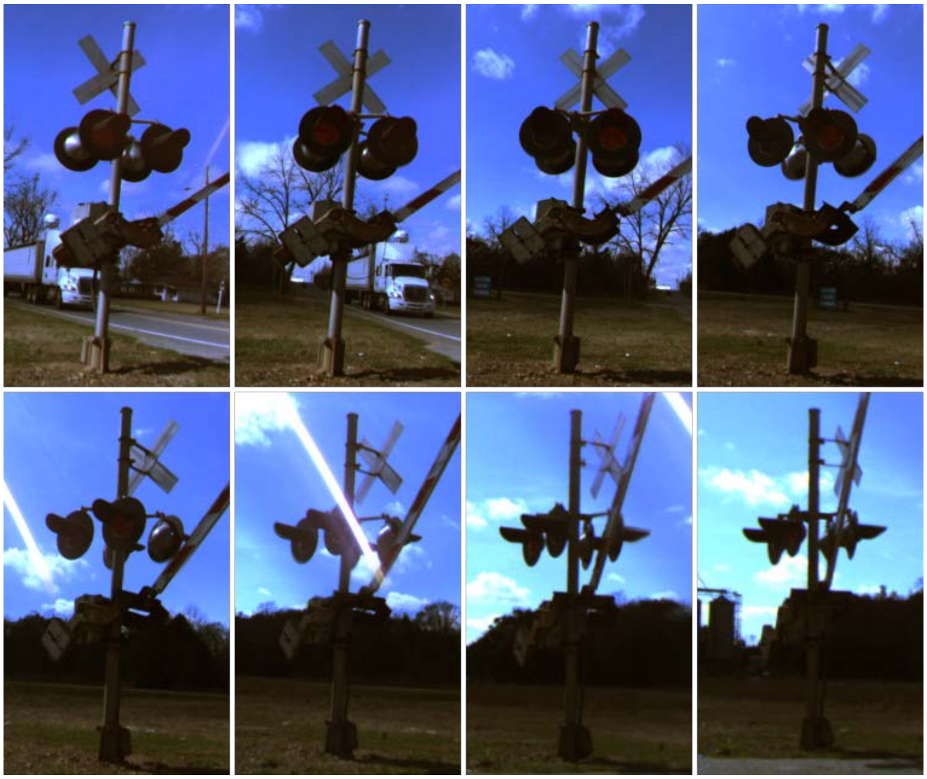


Figure 49. Object tracking images.

4.4 Video Zoom

This virtual environment serves as the main interface to the data system. It provides the user with a natural and intuitive way to navigate the large quantity of data. The user accesses higher resolution track data by simply zooming into the spherical video, which gives the user a sense of context. Figures 50 and 51 depict the effects of zooming into the virtual environment.

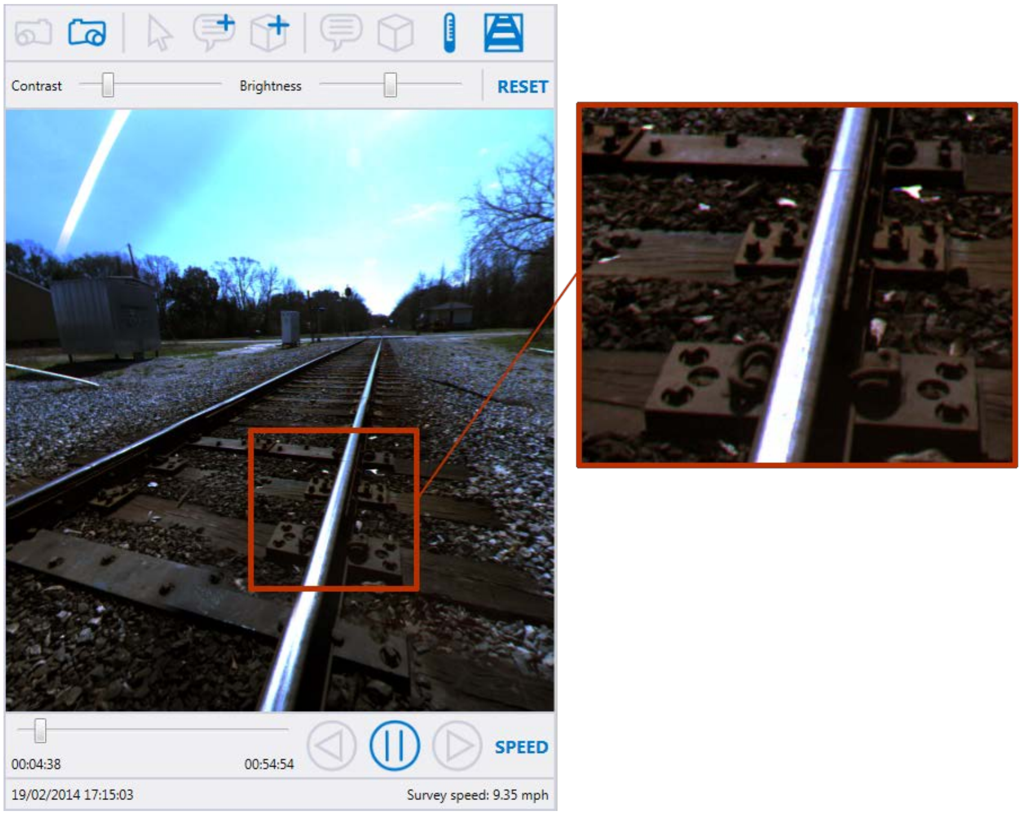


Figure 50. Detailed images through zooming.

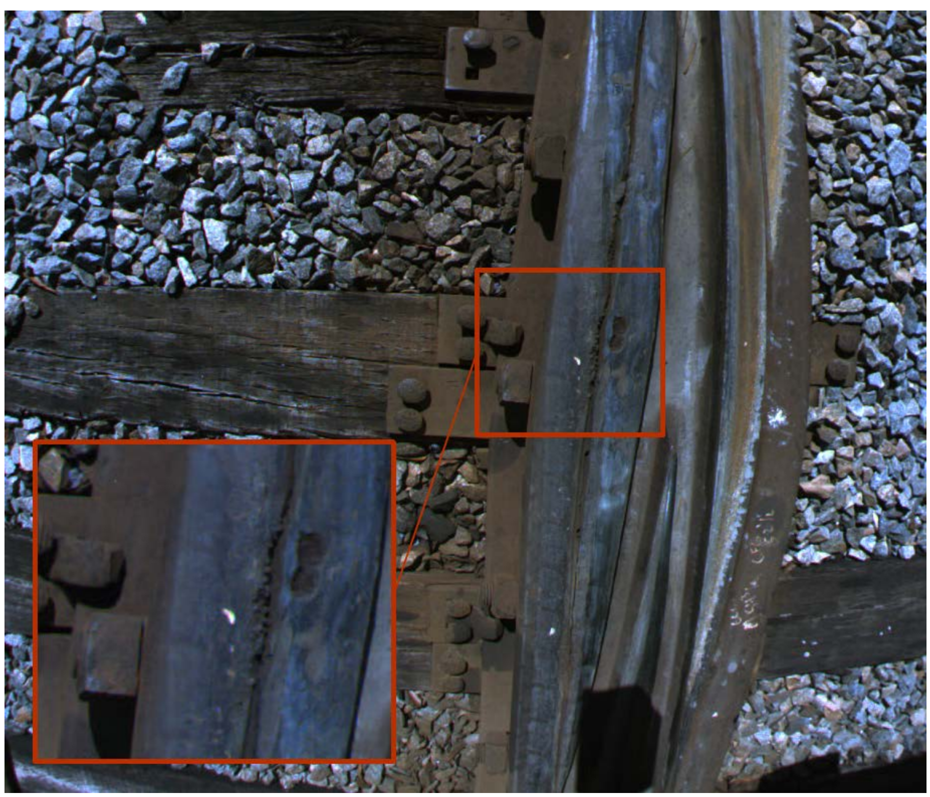


Figure 51. Fine details visible in zoomed image.

4.5 External Data Integration

The TTVS virtual environment is powerful because it not only displays data generated internally but also displays data imported from external sources; it binds together geo-located survey data from a number of survey processes and presents that data in a consistent and easy to understand manner. Gathering multiple survey data into one environment provides greater insight than is possible with the data in isolation and enables the user to quickly gain a comprehensive understanding of track condition. Figures 52 and 53 depict examples of external data imported into the TTVS virtual environment.

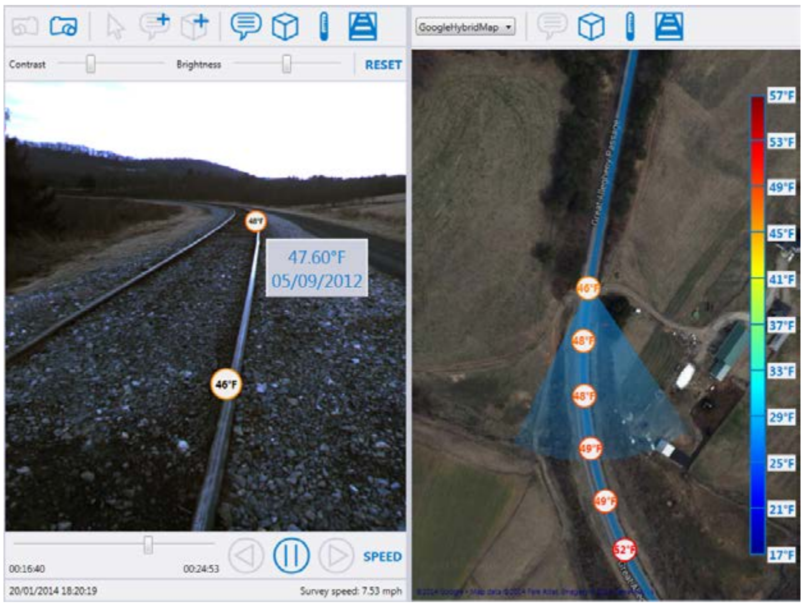


Figure 372. Rail temperature measurements imported into the TTVS virtual environment.

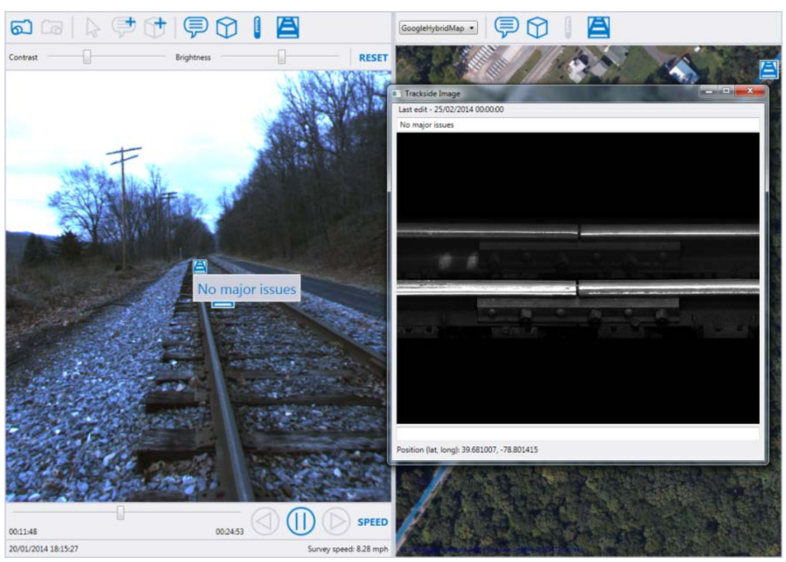


Figure 53. Joint bar images imported into the TTVS virtual environment.

5. Conclusions

Track maintenance survey processes currently require a significant amount of manual observation, which causes long and costly disruptions to normal track operation and exposes maintenance engineers to health and safety risks. The TTVS system developed for the FRA significantly reduces the time spent on track by capturing continuous high resolution video of the track and surrounding areas, and provided a complete suite of pre-processed visual data which enables the user to perform a virtual manual survey. Eventually, the system will enable operators to undertake surveys at track speed, enhance survey performance with new data, repeat surveys retrospectively, and view the history of an object or site by comparing new and old data.

The GUI developed to display TTVS data gives the user an easy-to-use interface to a large quantity of high-quality images which replicate the experience of actually being track-side as closely as possible, while adding new benefits, such as the ability to step back in time, or instantly leap ahead up the track. Additionally, the GUI demonstrates the effectiveness of an immersive video environment as a tool for visual surveys of the track and track-side, and as a data navigation framework used to visualize, interact with and create GIS data.

Trials of the TTVS system have successfully demonstrated an ability to attain detailed imagery of the track and the track-side at a resolution of 1 mm and at speeds of up to 28 mph (almost 150 percent design speed). Egomotion and object tracking algorithms performed well in the rail environment; this underlying technology enables the user to navigate video footage using a map interface as well as accurately create and manipulate GIS objects.

We believe that the use of the TTVS system will improve data quality, reduce survey costs, and lower the health and safety risks associated with track maintenance activities. The system captures a permanent record of the track condition, which allows engineers to repeat surveys retrospectively and in the safety of their own office. The prototype TTVS system tested on the WMSR and in the Panama City region of Florida has significant value as a survey system at track speeds of around 20mph. Section 6 contains recommendations for increasing the survey speed.

6. Recommendations for Further Development

The TTVS prototype's main limitation is the survey speed, which is dictated by camera resolution and frame rate. Using COTS spherical cameras to provide the desired level of detail limited the system to a maximum design speed of 20 mph. Survey speeds could be increased with a custom camera array utilizing sensors with higher resolution and higher frame rates. One potential sensor has been identified which would provide three times the frame rate and approximately a 100-fold increase in resolution.

By improving processing algorithms such as Egomotion, the geo-location of survey footage could be improved and with more precise geo-location and improved object tracking, automated inter-survey analysis using the TTVS system would become possible. An inter-survey analysis toolset could include facilities such as automatic change detection or the production of a timeline depicting the gradual change in a track-side feature.

This project has demonstrated the TTVS virtual environment's ability to bind together geo-located survey data from a number of survey processes and present it in a consistent and easy to understand manner. In the next phase in development, the project could start defining intuitive ways to visualize additional survey data, such as overlaying continuous data onto track imagery or displaying synchronized track geometry graphs alongside video playback. Further development should explore the ability to gather multiple survey data into one environment, which provides greater insight than is possible with the data in isolation.

If additional enhancements are made to the end user's toolset, the project could introduce new tools. These tools could include the capability to make measurements directly from video data, collate data on all GIS objects within a user-selected region, identify regions of track in need of further inspection, or provide a graphical indication of track health, based upon survey data imported into the virtual environment.

Finally, the technology developed by this project has valuable applications in other track inspection systems. For example, Createc understands that the minimum operating speeds of INS systems currently limit low speed track geometry measurements. An Egomotion system has no lower speed limit and could complement the INS component of an inspection system. Another project could explore this potential.

Abbreviations and Acronyms

BAYL	Bay Line
COTS	Commercial Off-The-Shelf
FRA	Federal Railway Administration
GIS	Geographic Information System
GPS	Global Positioning System
GUI	Graphical User Interface
INS	Inertial Navigation System
LIDAR	Light Detection and Ranging
PTZ	Pan, Tilt, and Zoom
TTVS	Track and Track-side Video Survey
WMSR	Western Maryland Scenic Railway