

Design of Dual-Road Transportable Portal Monitoring System for Visible Light and Gamma-Ray Imaging

Thomas P. Karnowski^{*a}, Mark F. Cunningham^b, James S. Goddard^a, Anil M. Cheriyyadat^a, Donald E. Hornback^a, Lorenzo Fabris^a, Ryan A. Kerekes^a, Klaus-Peter Ziock^a, E. Craig Bradley^a, J. Chesser^a, W. Marchant^c

^aOak Ridge National Laboratory, One Bethel Valley Road, Oak Ridge TN 37831

^bLawrence Livermore National Laboratory, Livermore CA 94550

^cSpace Sciences Laboratory, U. California-Berkeley

ABSTRACT

The use of radiation sensors as portal monitors is increasing due to heightened concerns over the smuggling of fissile material. Transportable systems that can detect significant quantities of fissile material that might be present in vehicular traffic are of particular interest, especially if they can be rapidly deployed to different locations. To serve this application, we have constructed a rapid-deployment portal monitor that uses visible-light and gamma-ray imaging to allow simultaneous monitoring of multiple lanes of traffic from the side of a roadway. The system operation uses machine vision methods on the visible-light images to detect vehicles as they enter and exit the field of view and to measure their position in each frame. The visible-light and gamma-ray cameras are synchronized which allows the gamma-ray imager to harvest gamma-ray data specific to each vehicle, integrating its radiation signature for the entire time that it is in the field of view. Thus our system creates vehicle-specific radiation signatures and avoids source confusion problems that plague non-imaging approaches to the same problem. Our current prototype instrument was designed for measurement of upto five lanes of freeway traffic with a pair of instruments, one on either side of the roadway. Stereoscopic cameras are used with a third “alignment” camera for motion compensation and are mounted on a 50’ deployable mast. In this paper we discuss the design considerations for the machine-vision system, the algorithms used for vehicle detection and position estimates, and the overall architecture of the system. We also discuss system calibration for rapid deployment. We conclude with notes on preliminary performance and deployment.

Keywords: Radiation monitoring, machine vision

1. INTRODUCTION

In this work we discuss the design considerations for a machine-vision system for detecting, tracking, and measuring the position of vehicles on a roadway with an application of gamma-ray imaging. The use of radiation sensors as portal monitors is increasing due to heightened concerns over the smuggling of fissile material and rapidly deployable systems that can detect significant quantities of fissile material in vehicular traffic are of particular interest. We have constructed a prototype, rapid-deployment portal monitor that uses visible-light and gamma-ray imaging to monitor multiple lanes of traffic from the side of a roadway^{1,2}. This system allows higher spatial selectivity because the gamma-ray imaging instrument can effectively “focus” on a chosen vehicle to measure its radiation signature, avoiding false positives that plague more conventional systems. There are several design considerations for this application that we discuss in detail, including: camera specifications, system calibration, and motion estimation accuracy.

The paper is divided into three main sections. In this first section we discuss the basic principle of operation of the instrument, focusing on how the machine-vision system can increase sensitivity of the gamma-imaging system. In the second section we discuss our machine-vision system at a high, functional level and describe the processing which is needed to accomplish our goal of real-time vehicle detection and measurement. The section concludes with a discussion of the main design considerations of the machine vision system, including the camera specifications, system placement for the roadside application, timing considerations, and stereo-vision requirements. Particular design considerations for

* T.P. Karnowski, karnowskitp@ornl.gov, 865-574 5732

processing, including: camera motion correction, and estimation; system calibration; and real-time bandwidth considerations unique to our dual roadside configuration are discussed in the next section. We conclude this section with some discussion of the additional design considerations for decreasing the computational load of the main image processing tasks. In our final section we discuss conclusions and future work for this unique instrument.

1.1 Principle of operation

While a detailed description of the imaging performance of these systems is beyond the scope of this paper (see other references^{1,2,3,4} for more details), for completeness we describe the principle of operation of the gamma ray imaging system and how machine vision can improve the sensitivity of these devices. Our instrument uses coded aperture imaging^{5,6,7} to image gamma-ray sources. This method is an indirect method of imaging in which a radiation-blocking mask casts direction-dependent “shadows” on position-sensitive detectors. Image reconstruction is conducted by cross-correlating the recorded data with the mask pattern. A one-dimensional imager can distinguish the position of sources in a direction parallel to the imager, but it cannot discern sources above or below the plane of the imager, and only provides minimal information on the range to the source (and then only at close distances).

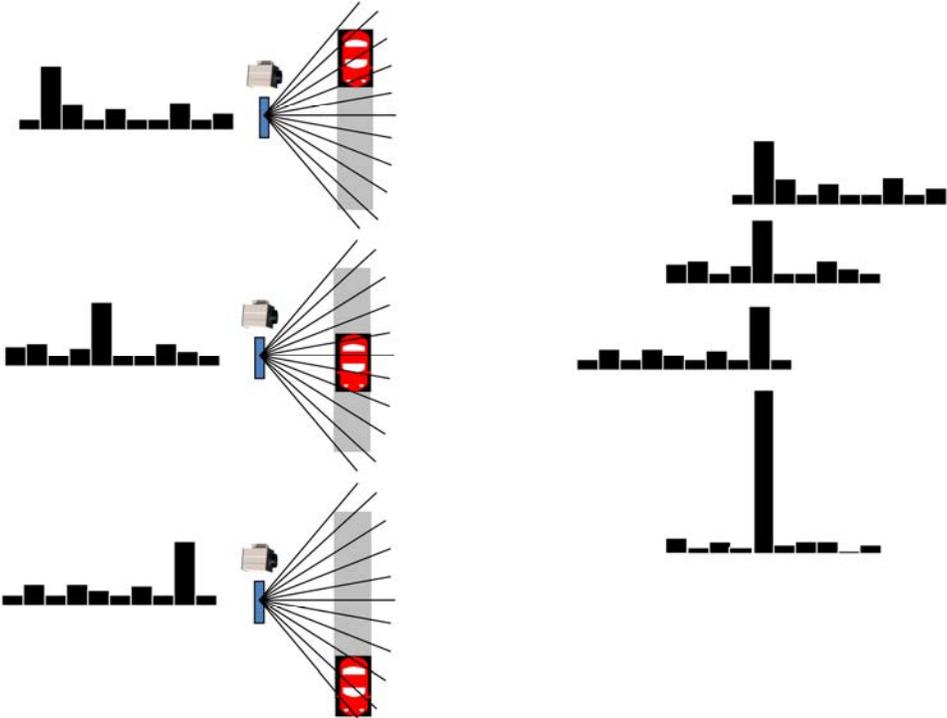


Figure 1. Conceptual illustration of visible-light and gamma-ray data fusion. Left: As a vehicle drives downward through the field of view of the gamma imager and machine vision system, the gamma imager forms a 1-D signal position-sensitive signal shown on the left side. The peak shifts with the vehicle from the top, middle, and bottom figures. Right: Since the machine vision system reports the vehicle position in each frame, we see that the gamma images can be shifted the correct amount and added to produce a stronger composite detection.

In the portable stand-off radiation detector imager⁸, gamma-ray images are mapped to real-world positions through knowledge of the instrument’s position (acquired by GPS, for example). For a stationary system, however, it is not possible in general to correctly shift and sum the gamma-ray images to integrate the data from a potential source transit. This is the purpose of the machine vision system in the Roadside Tracker as illustrated in Figure 1. Each gamma-ray imaging pixel (GRIP) is shown as constant-angle rays originating from the point to the left of the traffic lane. The machine vision system (depicted as the camera) detects vehicles in the field of view, and estimates their position in each frame. The position is used to estimate the times when the front of a vehicle crosses the gamma-ray pixel boundaries. Short exposure gamma-ray images are generated for the time that a vehicle resides in a single pixel. The sequential

images as the vehicle moves from one GRIP to another are shifted and added so that the vehicle is stationary in the summed image. Using this technique, the gamma-ray coded-aperture imager “harvests” the gamma-ray data specific to each vehicle, integrating its radiation signature for the entire time that it is in the field of view¹. In this fashion we are able to generate vehicle-specific radiation signatures and avoid source confusion problems that plague non-imaging radiation detection approaches to the same problem. In addition, we note that just as in the system of Ziock *et al*⁸, the motion of the vehicle allows us to estimate the range of sources through the parallax effect for better signal detections.

1.2 Machine vision system overview

The portable roadside tracker features two machine vision systems and gamma-ray imagers, one for each side of the roadway, to improve sensitivity and reduce the effects of occlusion. A diagram of the machine-vision systems is shown in Figure 2. The functional blocks consist of image acquisition, point detection and tracking, alignment estimation, and stereo reconstruction in a multi-threaded pipelined architecture running on Windows PCs. The gamma-imager and visible-light cameras are synchronized through a common clock triggering mechanism derived from a GPS time base to precisely synchronize data from both sides of the road. Three cameras are used in each system. The reference and stereo camera pair are used for the primary imaging with the reference camera serving as the “main” imager. A third camera, the alignment camera, images the roof of the trailer where fiducial marks are located. This third imager is used to estimate the position and orientation of the stereo-reference imagers relative to the gamma-imager coordinate system to improve measurement accuracy.

The second thread in the system pipeline performs feature detection and tracking on the images from the reference camera of the stereo pair using the OpenCV library⁹. The detected features are found using the “good features to track” algorithm¹⁰. The features are used to form “tracks” from frame-to-frame which are computed using the Lucas-Kanade optical flow algorithm¹¹. A parallel thread (the Alignment Estimate block) uses the images captured from the alignment camera to estimate motion and orientation changes in the camera system using fiducials on the roof of the trailer. The alignment camera position estimation, the detected points and tracks, and the stereo camera images are then processed in the next thread to perform stereo measurement of the 2D points through rectification and point matching. The resulting 3D points are then transformed from the reference camera coordinate system to “real world” coordinates relative to a point on the gamma-imager trailer.

After transfer to the “Master Computer” as shown in Figure 2, the track clustering block fuses the frame data from both sides of the road. This frame data includes the 3D coordinates of the tracks referenced to the local imager (the remote imager coordinates are transformed to the local point-of-view using a coordinate transformation estimated during a calibration phase as described later in this work). The 3D coordinates of the tracks are projected to the x-y plane (i.e., the road) and are clustered together to form vehicles. Our earlier implementation relied solely on monocular, frame-to-frame changes and followed a set of rules inspired by Beymer¹². Our latest implementation uses the “coherent motion regions” (CMR) algorithm¹³ adapted to operate on sets or “batches” of frames. The CMR algorithm clusters the tracks together by solving an optimization process subject to weights attached to the tracks which include relative velocity and spatial location.

The final steps are performed on a vehicle-by-vehicle basis after a detection has been made. Motion estimation consists of determining the position of the vehicle in each frame using the track information. Video event generation consists of determining the time each GRIP boundary is crossed; this information is sent to the gamma-imaging software via Ethernet links so that it can determine how to co-add the gamma images. Several candidate motion and position estimation methods have been explored for their accuracy¹⁴.

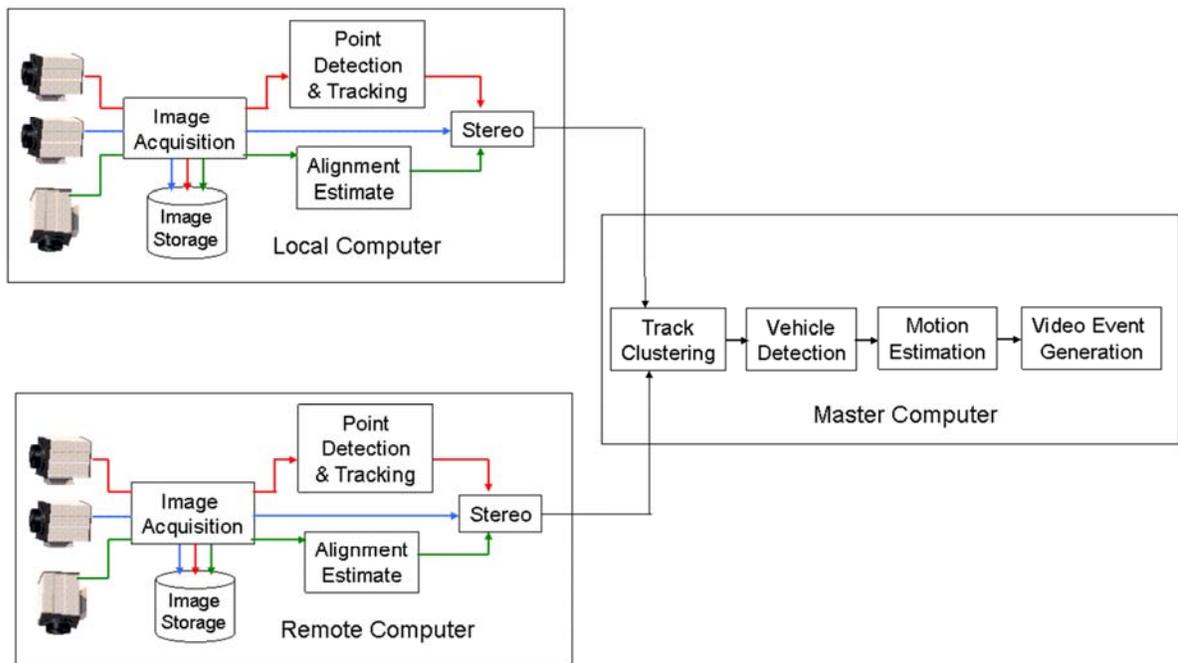


Figure 2. Illustration of machine vision system. Two essentially identical machine vision systems exist for both sides of the road, embodied in the “Local” and “Remote” computers. For each system the stereo/reference camera pair along with an alignment camera are depicted. After image acquisition and storage, the reference camera (red or upper trace) is used to perform point detection and tracking. In parallel the motion of the camera mast is estimated using the green or bottom trace which represents the alignment camera images. The stereo camera is then used along with the point detection and alignment estimate to create 3D estimates of the vehicle points. The final blocks perform track clustering, vehicle detection (identifying when a vehicle has entered and exited the field of view), and motion estimation

2. MACHINE VISION DESIGN CONSIDERATIONS

The machine vision system must locate vehicles and track them through the field of view of the gamma-ray imager. The overall goal is 100% vehicle detection and localization, enabling effective and accurate measurements of the gamma-ray signature can be performed. As in all engineering pursuits, there are tradeoffs between cost, risk, ease of implementation (related to cost), and commercial availability of components, all of which are important considerations. In this section we discuss the different functional blocks of the machine vision system and their design constraints. We conclude the section with an overview of computational considerations for the various processing blocks that comprise the remainder of our machine vision system.

2.1 Image Acquisition

Image acquisition considerations include the camera location, specifications, and constraints. We address these considerations in separate sub-sections that pertain to the cameras, field-of-view, resolution requirements, timing issues, and camera placement.

Mast height

Avoiding occlusion is a major issue in machine vision applications that deal with traffic. We chose to address this problem by increasing the height of the camera as much as possible to minimize occlusion. In addition, because our system is intended to be mobile, we assumed we could not rely on fixed camera installations. Therefore we elected to elevate our cameras using a commercially-available deployable mast. Our initial design used a 24' mast constructed by Hi-Pod¹⁵. For our current system, we assumed five lanes of traffic, each 3.5 m wide, with a mast distance from the road

of 5 m. However, since there are matching systems on both sides of the road we assume that we need only consider the occlusion in the nearest three lanes (assuming a five-lane roadway). Four different vehicle types were studied: a large truck or 18-wheeler, an SUV, a midsize sedan, and a compact car. To estimate the needed mast height, we used a simple geometrical model to estimate the occlusion caused by different sized vehicles at different mast heights. This is illustrated in Figure 3. A ray is constructed from the camera atop the mast to the roof of the occluding vehicle. There are four thresholds of interest. The first is complete occlusion and occurs when the occluded vehicle height is less than H_s placing it below the ray. At the other extreme, if the vehicle is sufficiently far away (D_{all}) no occlusion occurs. The other two cases are when some fraction of the roof is visible but none of the side is visible, and when the entire roof is visible but only a portion of the side is visible. The equations of interest are easily derived from similar triangles and are shown below:

$$D_{all} = \frac{D_1 H_m}{H_m - H_1} \quad (1)$$

$$H_s = (D_{all} - D_2) \left[\frac{H_m - H_1}{D_1} \right] \quad (2)$$

$$H_R = (D_{all} - D_2 - W_2) \left[\frac{H_m - H_1}{D_1} \right] \quad (3)$$

$$F_R = \left[\frac{1}{H_R - H_S} \right] H_2 - \frac{H_S}{H_R - H_S}, H_S < H_2 < H_R \quad (4)$$

$$F_S = \frac{H_2 - H_R}{H_2}, H_2 > H_R \quad (5)$$

There are some additional simplifications that can be made. Since there are imagers on both sides of the road, we assume that we are only concerned with occluding vehicles in lanes 1 and 2, with occluded vehicles in lanes 2 and 3. The analysis was computed by establishing a mast height, then iterating through all combinations of vehicles in lane 1 with occluded vehicles in lanes 2 and 3, then repeating for vehicles in lane 2 with occluded vehicles in lane 3. The final results were tallied and cases where no occlusion is found are omitted. These results are tabulated in Table 1 below for cases where some obstruction occurs. Mast heights of 15.25m to 38m were evaluated. Essentially at mast heights of 38m or more there is no occlusion in all cases, so this is not shown in the table. We note that these estimates assume that vehicles travel in the middle of the lanes. Ultimately, we elected to choose a 15.25m mast (50') due to the availability of a commercial unit¹⁶ at a reasonable cost; this represents a logical "next step" from our preliminary work with the 8m mast.

Camera

The main considerations for the cameras included interface, resolution, acquisition speed, power, and time synchronization capabilities. Our previous prototype utilized a FireWire camera. The limitations of the FireWire format included the limited cable length, but this was easily circumvented by using an amplification system. In our application the time synchronization properties proved to be the most limiting factor in a camera choice. First, we required that all visible-light images be acquired simultaneously for accurately computing stereo data while correcting for the mast motion and synchronizing results from the two sides of the roadway. Second it is crucial that the gamma-ray and video imaging systems use a common time base in order to accurately co-add the gamma radiation data. Finally to verify timing we also desired an absolute time stamp that could be embedded within the camera image stream. The objective here is to detect dropped frames in either the camera buffer or the operating system, since undetected dropped frames would prevent correct synchronization of timing between the gamma imaging system and the video system. After a detailed study and search, we were able to identify a camera vendor with a camera line that featured the ability to time-tag every frame at the camera using the cameras' onboard Ethernet clock. To ensure synchronization of the Ethernet clocks of the different cameras, we made use of an option that allowed resetting this clock externally. Specifically, we

reset the clocks every second using a reference pulse synchronized to universal time obtained from a GPS timebase. With this approach we can ensure camera frame synchronization and absolute time, provided that we do not lose more than one second of frames. Our camera resolution is based on the heuristics of our previous system which yielded good performance with respect to vehicle visibility and point detection using a camera at 8 m height with 640 x 480 pixels. Since our newer design features a taller mast, we experimentally verified that we would get comparable point coverage at the new height with a larger format camera. Consequently the Gevicam model GP-3780C was selected, which uses a 1032 x 779 pixel, 8.47-mm CMOS imager¹⁷.

Table 1. Occlusion for different mast heights and vehicles in different lanes. TT=tractor-trailer; SUV=sport utility vehicle. The fraction (Frct) of the roof and side at the different mast heights are shown.

| Vehicle1 | Lane1 | Vehicle2 | Lane2 | Roof Frct at 15.25m | Side Frct at 15.25m | Roof Frct at 23m | Side Frct at 23m | Roof Frct at 30.5m | Side Frct at 30.5m |
|----------|-------|----------|-------|---------------------|---------------------|------------------|------------------|--------------------|--------------------|
| TT | 1 | TT | 2 | 1.00 | 0.65 | 1.00 | 1.00 | 1.00 | 1.00 |
| TT | 1 | SUV | 2 | 0.63 | 0.00 | 0.99 | 0.00 | 1.00 | 1.00 |
| TT | 1 | Sedan | 2 | 0.48 | 0.00 | 0.91 | 0.00 | 1.00 | 1.00 |
| TT | 1 | Compact | 2 | 0.31 | 0.00 | 0.85 | 0.00 | 1.00 | 1.00 |
| TT | 2 | TT | 3 | 1.00 | 0.46 | 1.00 | 0.78 | 1.00 | 1.00 |
| TT | 2 | SUV | 3 | 0.25 | 0.00 | 0.77 | 0.00 | 0.99 | 0.00 |
| TT | 2 | Sedan | 3 | 0.03 | 0.00 | 0.64 | 0.00 | 0.90 | 0.00 |
| TT | 2 | Compact | 3 | 0.00 | 0.00 | 0.52 | 0.00 | 0.84 | 0.00 |

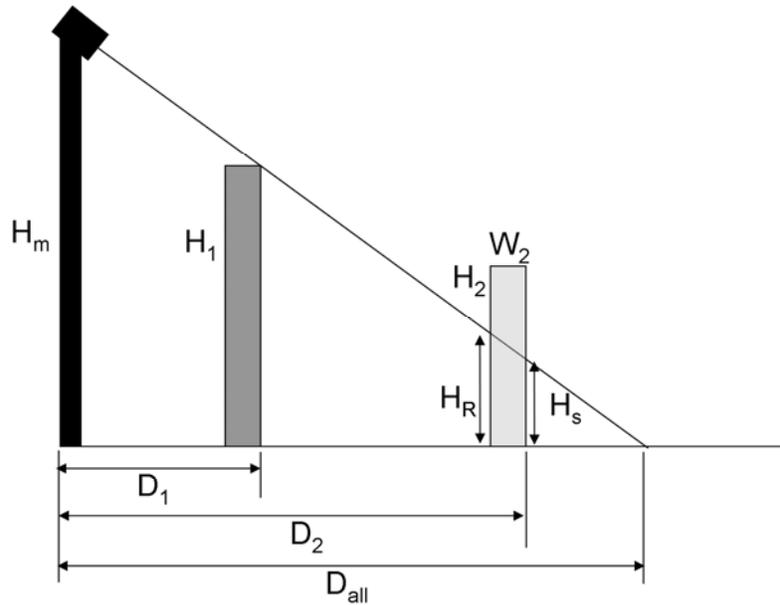


Figure 3. Simple geometric model for occlusion study to dictate camera placement and mast height. The mast height is H_m and the occluding vehicle has a height of H_1 at a distance of D_1 . Any vehicles at a distance of D_{all} or greater will not be occluded. Vehicles at a height of H_s or less at distance D_2 will be completely occluded. Between those thresholds, the camera should see a fraction of the roof for vehicles whose height is between H_s and H_R . Some fraction of the side can be seen for vehicles whose height is greater than H_R .

Field of view

We selected the field of view (FOV) for the visible-light imaging system to meet or exceed that of the gamma imager. The gamma imager FOV, in turn, is a function of the focal length and mask pattern of the instrument¹. In addition, we

wanted our visible-light resolution to at least match that of our previous system since it gave good performance in finding features on automobiles. In that system, the resolution was roughly 35 mm / pixel at 5 m range, measured by simply analyzing images from that system. For our design we use the following equations:

$$F_{ov} = \frac{HP_x}{F} \sqrt{H_m^2 + R^2} \quad (6)$$

$$R_s = \frac{F_{ov}}{H} = \frac{P_x}{F} \sqrt{H_m^2 + R^2} \quad (7)$$

where F_{ov} is the field of view, H is the number of horizontal pixels, P_x is the size of the pixels, F is the lens focal length, H_m is the mast height, R is the range to the target, and R_s is the resolution in pixels per distance. We tabulate these values for three different focal length lenses and five different traffic lanes in Table 2. From these results we see that all focal lengths give sufficient resolution, but the 5-mm focal length does not cover the gamma imager field of view at the longest ranges. The 3-mm and 4-mm lenses cover the field of view adequately. For our system, off-the-shelf fixed lenses of 3.5 mm focal length were chosen for the reference and stereo cameras.

Table 2. Field-of-view for different lens sets. The range is assumed to be the center of the lane of interest. The visible light camera is 15.25 m high and the imager is a 1032x779 camera with a sensor size of 8.47 mm and 4.65 micron pixels

| Lane (Range) | Gamma Imager FOV (m) | Visible Light System | | | | | |
|--------------|----------------------|----------------------|---------------------|---------|---------------------|---------|---------------------|
| | | 3.0 mm | | 4.0 mm | | 5.0 mm | |
| | | FOV (m) | Resolution (pix/mm) | FOV (m) | Resolution (pix/mm) | FOV (m) | Resolution (pix/mm) |
| 1 (6.75 m) | 9.8 | 26.5 | 3.9 | 19.9 | 5.3 | 15.9 | 6.5 |
| 2 (10.25 m) | 14.9 | 29.2 | 3.5 | 21.9 | 4.7 | 17.5 | 5.9 |
| 3 (13.75 m) | 20.0 | 32.7 | 3.2 | 24.5 | 4.2 | 19.6 | 5.3 |
| 4 (17.25 m) | 25.1 | 36.7 | 2.8 | 27.5 | 3.7 | 22.0 | 4.7 |
| 5 (20.75 m) | 30.1 | 41.1 | 2.5 | 30.8 | 3.4 | 24.6 | 4.2 |

Timing

The video system and gamma imaging system must maintain a common time base to ensure the best quality gamma imaging (and consequently the highest signal-to-noise ratio). We also required that both systems (local and remote) start simultaneously and synchronously so that the frame number could be used to verify synchronization. Some methods we considered included wireless triggering and optical line-of-sight triggering, but instead we elected to use a GPS-based timing system. Two precision GPS time bases²² were acquired and customized by the manufacturer to provide two trigger pulses in sync with the GPS universal time clock (UTC), one at one-second intervals and a second at a selectable interval. The former was used to reset the internal camera clocks as described above. The latter was used to trigger the camera shutter releases. The leading edge of all of the pulses are aligned to better than a microsecond to ensure a fixed phase relationship for the trigger signals. In addition to releasing the camera shutters, the faster pulses are detected by the gamma-ray data system and entrained in the list mode gamma-ray data. This is sent to the host computer in one-millisecond frame times so that each gamma-ray event is known to a millisecond. At one second intervals, an embedded processor²³ computer that uses the GPS as a time server also sends the absolute time to the gamma-ray system. This is also entrained in the data stream so that the absolute time of the shutter release pulses can be verified. When the computers on both sides of the road are configured for a data acquisition run, the embedded processors on both sides of the road use the NTP server in the GPS time base to gate the 1-Hz and shutter release pulses to start on precise 15 second UTC intervals (on the minute, and 15, 30, and 45 seconds afterward). This delayed start is long enough that any house-keeping performed by the different computers in preparation of a run can be completed. Precise stopping of the

acquisition is much less important and thus it is possible for both sides of the road to actually acquire different amounts of data, but simultaneous starting is ensured by our procedure.

The tolerance of the timing delay between both sides of the road was estimated based on maximum vehicle speed and blurring effects. Ideally, if we take an image with one camera, we want the other camera image to be taken before the vehicle has traveled some number of pixels. The design equation of interest is

$$S = \frac{kF_{ov}}{VH}, \quad (8)$$

where k is the maximum pixel shift, F_{ov} is the field of view in meters, V is the velocity of the vehicle in m/s, and H is the number of horizontal pixels in the camera field of view. With a maximum speed of 80 MPH and a pixel shift of $k=0.10$ pixels, the chosen camera and lens indicate the time lag between images must be less than $70 \mu\text{s}$. This is easily met by the $\pm 300 \text{ ns}$ precision between the shutter release pulses on the two sides of the roadway.

Our camera frame rate was ultimately determined by our ability to write the images to disk. For debugging and diagnostic purposes we elected to save all data to disk in uncompressed format. The camera provides Bayer image data such that each image uses 8 bits per pixel. We found experimentally that we could write data to disk at a rate of 20 fps with no dropped frames for periods up to 30 minutes in duration which was regarded as sufficient for development purposes. Higher rates could be accommodated through compression (at some computational cost for the compression algorithm) or for limited amounts of time, however we considered this to be a reasonable starting point and a good step above our previous development system which ran at 15 fps. For completeness, although the overall frame rate may well be limited by various other factors, particularly processing speed of steps in the pipeline, we estimated the desired rates based on vehicle speeds and time in field of view. A table of these values is shown below. Clearly a wider field-of-view allows more video frames for monitoring any single vehicle of interest.

Table 3. Video frames for an 80 MPH vehicle traversing the field of view in the nearest lane (6.75m range) at different frame rates with the chosen mast and camera.

| Frame rate (fps) | 3.0 mm focal length | 4.0 mm focal length | 5.0 mm focal length |
|------------------|---------------------|---------------------|---------------------|
| 10 | 7.4 frames | 5.6 frames | 4.4 frames |
| 15 | 11.1 frames | 8.35 frames | 6.7 frames |
| 20 | 14.8 frames | 11.1 frames | 8.9 frames |
| 25 | 18.5 frames | 13.9 frames | 11.1 frames |
| 30 | 22.2 frames | 16.7 frames | 13.3 frames |

Stereo Rig

Our initial system used a monocular camera to estimate vehicle position. This was done using a homography transform which roughly mapped the image plane of the sensor to the real-world through a calibration process. For this phase of development we chose to obtain range information through the use of stereo imaging for a number of reasons. First, the range imaging helps determine the relative position of each imaging unit allowing us to establish a single coordinate system for the full system which is useful for fusing multiple views from both sides of the road. Second, the depth information enhances the separation of vehicles in neighboring lanes. Finally, the estimate of vehicle height benefits both the segmentation process, helps to detect partial occlusions, and also allows us to predict when full occlusions might occur which may allow us to form gamma-ray images of vehicles that are completely hidden from view of the machine vision system.

The accuracy of the stereo depth perception is an important design factor as this dictates the size of the stereo baseline required. We use the following equation for stereo depth accuracy²¹

$$\Delta R = \frac{P_x R^2}{FB} \Delta d, \quad (9)$$

where P_x is the pixel size, R is the range, F is the focal length, B is the baseline (distance between cameras), and Δd is the disparity in pixels. Our initial assumption was that we could achieve a subpixel accuracy of $\Delta d = \frac{1}{4}$ pixel. The other parameters are already limited as described above. To ensure that the optical system design did not affect the gamma-ray imaging, we chose a depth resolution at roughly 1/10 GRIP. This results in the design performance given in table 4 below for stereo baselines of 2500 mm, 2000 mm, 1000 mm, and 750 mm. A 2-m baseline achieves the desired goal of roughly 1/10 the depth resolution.

Table 4. Stereo depth resolution as a function of range for different stereo baselines with our specified camera.

| Lane | Gamma Range (m) | Gamma FOV (m) | GRIP Size (cm) | Video Range (m) | Delta R 2.5 m | Delta R 2.0 m | Delta R 1.0 m | Delta R 0.75 m |
|------|-----------------|---------------|----------------|-----------------|---------------|---------------|---------------|----------------|
| 1 | 6.75 | 9.8 | 42.6 | 17.4 | 4.02 | 5.03 | 10.06 | 13.41 |
| 2 | 10.25 | 14.9 | 64.8 | 19 | 4.8 | 6 | 11.99 | 15.99 |
| 3 | 13.75 | 20 | 87 | 21.1 | 5.91 | 7.39 | 14.79 | 19.72 |
| 4 | 17.25 | 25.1 | 109.1 | 23.5 | 7.34 | 9.17 | 18.34 | 24.46 |
| 5 | 20.75 | 30.2 | 131.3 | 26.2 | 9.12 | 11.4 | 22.8 | 30.4 |

A smaller baseline may be used, provided better subpixel accuracy can be achieved, but for this phase of development we elected to use a conservative subpixel resolution. Our custom stereo rig is shown in Fig. 4. Our design was limited by the maximum mast payload of 18.1 kg. The resulting rig is based on an aluminum pipe in order to accommodate requirements on thermal expansion and weight. The cameras are mounted on kinematic mounts and are covered with a shield to block excess sunlight (not shown in Figure 4).



Figure 4. Photograph of stereo rig with 2000-mm baseline. The reference and stereo pair (A and B) are on the extreme ends and the alignment camera (C) is seen pointing downward to the right of the signal distribution box in the center of the unit.

2.2 Mast Motion Correction Design

In our system potential movement of the mast requires special consideration. Since we cannot assume that the mast will be stationary as traffic and atmospheric conditions change, we must use methods to estimate the mast position so we can correct for these changes and provide accurate measurements of tracked vehicle positions. We assume that the stereo rig is a rigid unit therefore we need only estimate mast pose changes (rotation, elevation, and translation) for the entire unit as a whole. This is accomplished with the use of a third “alignment” camera which looks down from the mast at the roof of the trailer where we have mounted fiducial marks (Figure 5). Since the required field of view of this camera is much smaller than the traffic-viewing cameras, a much longer focal length is selected (12.5 mm) which produces an estimated resolution at full mast deployment of 5 mm / pixel for measurement accuracy of 1 mm/pixel lateral, 6 mm vertically, and approximately 0.1 degree of rotation sensitivity. Fiducial marks are affixed in the trailer corners and deliver high spatial-frequency content in a small view, with rotation invariance; they serve as excellent targets for normalized correlation operations¹⁸. In addition a set of corner points are also positioned on the roof. Finally the roofs have been

painted with flat gray paint to reduce sunlight reflections that can be seen from the opposite mast at certain times of day. The mast correction procedure consists of taking the first frame of video and locating the bulls-eye targets by scanning the entire image. This process is computationally intensive and cannot be performed in real-time on every frame. Hence, for subsequent frames we use a much smaller search window since the position of the image is not expected to shift significantly from frame to frame. As a result, the first frame takes roughly 80 ms (or about 1.6 video frames) but subsequent searches only require approximately 2 ms per frame, thus the correction process catches up almost immediately and proceeds in real-time. The system was designed to include guy wires to stabilize the mast, but these have proven unnecessary in our initial testing. We plan to further analyze their utility and necessity in future deployments.



Figure 5. Image of the roof of one of the trailers from the alignment camera showing the calibration targets.

2.3 System Calibration and Setup

To obtain sensitive gamma-ray detections, maintaining the calibration between the video and gamma-ray imagers is crucial. A complete mapping is done by taking a series of simultaneous gamma-ray and visible-light images of a radiation source at a number of locations with the system. If we assume for the moment that the visible-light images are taken without motion of the visible-light cameras between these different exposures, we have established a transformation between the gamma-ray and visible-light camera fields of view. Because this calibration procedure is impractical when the unit is deployed next to a busy roadway, we take advantage of the fact that the field of view of the gamma-ray imager is fixed with respect to the trailer. We can therefore use the location of the visible-light targets on the roof of the trailer to estimate the gamma-ray imager field of view, once this calibration has occurred. However, for this procedure to work, we need to compensate for the motion of the visible light cameras with respect to the trailer. There are three primary causes of such motions: mast height repeatability (relatively minor), repeatability of guy wire placement, and motion due to vibration and wind. In practice, the video system works in real world coordinates and not those of the gamma-ray imager. These coordinate systems are established quickly at the start of a system deployment. Overall there are three aspects of calibration: intrinsic camera calibration, extrinsic camera calibration, and calibration between the two systems on opposite sides of the roadway. In addition, we require the specification of lane boundaries and an image mask to decrease the computational load on point detection and stereo rectification.

The camera calibration is done off-line using the Callab Calibration Toolkit from DLR Calibration Laboratory¹⁹. First we perform an intrinsic camera calibration which is done only once per camera using a standard chessboard target. The extrinsic camera calibration is more complex and requires the specification of two transforms: (1) the transform between the stereo and reference cameras for stereo vision reconstruction, and (2) the transform between reference and alignment camera for position estimation. Both procedures use a chessboard target as well and require that stereo rig is mounted on the mast. A sequence of 12 different views of the checkerboard target are presented at various poses and mast heights. From these, the camera calibration toolkit is used to estimate the transformations between the stereo-reference camera pair and the reference-alignment camera pair.

The calibration between the two systems on the opposite sides of the roadway and other deployment-specific calibration steps are conducted with a customized software tool. The tool is run on the “Master” computer and requires that the video computers run in data acquisition mode with the triggering signal available. However, it is not important that the timing of the side-to-side acquisition be precise since the relative positions are calculated on-the-fly for both systems. The calibration procedure begins with the acquisition of video signals from both sides of the road. Once a set of images are obtained from all six cameras, several points of interest are identified in the local reference image and their corresponding matches are found in the local stereo, remote reference, and remote stereo image. The points are converted from image pixels to stereo matches and then to real-world coordinates relative to each trailer. Any motion in the mast when the images were taken is removed using the mast motion correction procedure so that these points reflect the real-world position of the selected features. From there, the points are used to estimate a least-squares error transformation between the two trailer coordinate system. The transformation (a 3x4 matrix) is stored for all runs until the trailers are moved again.

The next step involves the selection of the traffic lanes. This is a very simple step and could possibly be automated although operator oversight would be desired. The lanes are selected in the Local view (defined as the trailer that houses the Master computer) and their positions transformed to the viewpoint of the other (Remote) trailer. The lane specification is used largely for information purposes in the current system design to help identify the vehicle location.

The final step is the selection of an image mask. This is used to reduce the computational load of the point selection algorithm and the stereo rectification process. The mask is chosen to remove areas where vehicles should not appear. We note that we can track vehicles that travel into these regions but we do not compute new points for segmentation in these areas. Furthermore the rectification process only warps regions outside the mask area to minimize the processing time.

2.4 Data Flow Bandwidth Considerations

A final special consideration is the data flow and bandwidth requirements. Each machine vision system performs compute-intensive operations “locally”. A Master computer is used to merge the results from both sides of the road. It receives the information from both systems via an Ethernet link. Communications must respect the bandwidth limitations of this link, which is particularly important for the remote trailer since we are using wireless bridge to span the roadway. This means that the full visible-light images cannot be sent to the Master computer. Since the segmentation information is generated before this link, the video frames are needed at the Master only for display purposes. Thus we elected to send a compressed copy of all the frames. The compression must be sufficient to reduce the bandwidth of the data below the limitations of the wireless link but must retain enough visual acuity to allow reasonable visualization of the vehicle in question. We also elected to use color images to aid the visualization. The pure bandwidth required for 24-bit color images at our resolution and frame rate is 368 MBit/s which is well over the theoretical limit of the wireless system (100 MBit/s). Using the images in their native Bayer format would reduce this by a factor of 3 to about 123 MBit/s but this is still too costly as the segmentation and gamma-ray data must also pass through this link. We thus chose to use JPEG compression using software libraries already available in the .NET programming environment. Compression using a quality factor of 20 required an acceptable 20 ms per frame and meets our visualization and bandwidth requirements. For speed considerations we take the Bayer image, shrink it by a factor of 3, then convert to color and compress. The resulting images are sufficient for video display passing 20 frames per second and allow the system to operate in real-time.

The net bandwidth required is allocated as follows. We assume 16 vehicles / second with roughly 20 selected feature points per vehicle. Each point has eight 4-byte values associated with it (image plane and real-world coordinates along with an index into the previous frame for building tracks) for a total of 32 bits / point or 640 bytes/vehicle. For 16 total vehicles this is 10,240 bytes. Adding a compressed frame of video at roughly 18,000 bytes produces a total of just under 30,000 bytes / frame; thus at 20 fps we have a requirement of under 5 MBit/s. As a sanity check, a general purpose traffic lane receives a maximum flow estimated at 2000 vehicles per hour or about 0.5 vehicles / second at 60 MPH²⁰. Extrapolating to 80 MPH and 5 lanes of traffic increases this to roughly 4 vehicles / second, thus we believe our estimate of 16 vehicles / second is fairly conservative so long as traffic is moving freely.

3. CONCLUSIONS

In this work we discussed different design aspects of a machine vision system for use with a coded-aperture gamma-ray imager. The main design components of camera placement and specification, along with motion correction and

calibration were discussed. Our system has been tested in simple roadside settings on site at ORNL and will be tested under more stringent conditions in the future. Our processing pipeline was discussed at a high level and special attention was given to the data flow rate for the system. For future work we plan to further test the system; quantify the relationship between range, accuracy of tracking, and gamma signature quality; and improve the point detection and tracking by improving the track clustering and point clustering algorithms.

4. ACKNOWLEDGMENTS

We wish to recognize the mechanical support of Curtis Fitzgerald in developing the system hardware. We acknowledge the funding support of the Department of Homeland Security Domestic Nuclear Detection Office. This paper was prepared by Oak Ridge National Laboratory, Oak Ridge, TN, operated by UT-BATTELLE, LLC for the U.S. Department of Energy under Contract DE-AC05-00OR22725.

REFERENCES

- [1] Ziock, K.P., et. al., "The use of gamma-ray imaging to improve portal monitor performance", IEEE Transactions on Nuclear Science 55(6), 3654-3664 (2008)
- [2] Ziock, K.P., et. al., "Performance of the Roadside Tracker Portalless Portal Monitor", presented at Proc. IEEE Conf. on Nuclear Sciences, paper in preparation (2009)
- [3] Fenimore, E.E., "Coded aperture imaging: predicted performance of uniformly redundant arrays", Applied Optics 17(22), 3562-3570 (1983)
- [4] Ziock, K.P., et al., "Large area imaging for long-range, passive detection of fissile material", IEEE Transactions on Nuclear Science 51, 2238-2244 (2004)
- [5] R. H. Dicke, "Scatter-Hole Cameras for X-Rays and Gamma Rays", Astrophys. J. 1(153), L101 (1968).
- [6] J. G. Ables, "Fourier Transform Photography: A New Method for X-ray Astronomy", Proc. Astron. Soc. Aust. 1(1), 172 (1968).
- [7] Carolil, E. et al., "Coded aperture imaging in X- and gamma-ray astronomy", Space Sci. Rev. 45, 349-403 (1987)
- [8] Ziock, K.P., et. al., "A fieldable-prototype, large-area, gamma-ray imager for orphan source search", IEEE Transactions on Nuclear Science 55(6), 3643 - 3653 (2008)
- [9] <http://sourceforge.net/projects/opencvlibrary/>
- [10] Tomasi, C. and Shi, J., "Good features to track", Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 593-600 (1994)
- [11] Lucas, B., and Kanade, T., "An iterative image registration technique with an application to stereo vision", Proc. DARPA Image Understanding Workshop, 121-130 (1981)
- [12] Beymer, D. et al, "A real-time computer vision system for measuring traffic parameters," Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 495-501 (1997).
- [13] Cheriyyadat, A.M., Bhaduri, B.L. and Radke, R.J., "Detecting multiple moving objects in crowded environments with coherent motion regions", Proc. Sixth IEEE Computer Society Workshop on Perceptual Organization in Computer Vision (POCV) , 1-8 (2008)
- [14] Karnowski, T.P., et.al., "Motion estimation accuracy for visible-light/gamma-ray imaging fusion for portable portal monitoring", Proc. SPIE 7538 (2010)
- [15] <http://www.hi-pod.com/>
- [16] <http://www.floatograph.com/>
- [17] <http://www.gevicam.com/>
- [18] Gonzalez, R.C. and R. E. Woods, *Digital Image Processing*. Addison-Wesley (1992)
- [19] <http://www.robotic.dlr.de/CalLab>
- [20] Kwon, J. and Varaiya, P., "Effectiveness of California's High Occupancy Vehicle (HOV) system", Transportation Research Part C: Emerging Technologies 16, 98-115 (2008)
- [21] Ma, Y., Soatto, S., Kosecka, J. and Shankar Sastry, S., *An Invitation to 3-D Vision*, Springer (2003)
- [22] http://ptfinc.com/dsheets_2008/3203A_04A_08.pdf
- [23] <http://www.inhand.com/products/legacy-products>