Simulation and modeling of return waveforms from a ladar beam footprint in USU LadarSIM[1]

Scott Budge, Brad Leishman and Robert Pack
Center for Advanced Imaging Ladar
Utah State University, Logan, UT 84322-4170

ABSTRACT

Ladar systems are an emerging technology with applications in many fields. Consequently, simulations for these systems have become a valuable tool in the improvement of existing systems and the development of new ones. This paper discusses the theory and issues involved in reliably modeling the return waveform of a ladar beam footprint in the Utah State University LadarSIM simulation software. Emphasis is placed on modeling system-level effects that allow an investigation of engineering tradeoffs in preliminary designs, and validation of behaviors in fabricated designs. Efforts have been made to decrease the necessary computation time while still maintaining a usable model.

A full waveform simulation is implemented that models optical signals received on detector followed by electronic signals and discriminators commonly encountered in contemporary direct-detection ladar systems. Waveforms are modeled using a novel hexagonal sampling process applied across the ladar beam footprint. Each sample is weighted using a Gaussian spatial profile for a well formed laser footprint. Model fidelity is also improved by using a bidirectional reflectance distribution function (BRDF) for target reflectance. Once photons are converted to electrons, waveform processing is used to detect first, last or multiple return pulses. The detection methods discussed in this paper are a threshold detection method, a constant fraction method, and a derivative zero-crossing method. Various detection phenomena, such as range error, walk error, drop outs and false alarms, can be studied using these detection methods.

KEYWORDS: laser sensors and systems, waveform modeling and processing, LADAR, simulation

1. INTRODUCTION

LadarSIM is a Matlab and Simulink-based LADAR system simulator designed and developed by the Center for Advanced Imaging LADAR (CAIL) at Utah State University. It is a tool for system analysis, error source modeling, and specifically assists in the design and development of new LADAR systems. The appeal of LadarSIM lies not only in its usefulness but also in its speed and user-friendly software design.

In previous versions of LadarSIM, assumptions and generalizations were made in order to decrease the computation time and to increase the speed of the simulation. Most notably among these, was the use of a simple rectangular pulse to model the emitted and returned energy received at the detector. For flat surface returns this simple model is sufficient, however, when a ladar beam is reflected off of a multi-faceted surface, the simple model becomes deficient. In order to analyze and understand the various radiometric phenomena associated with actual ladar systems, improvements have been made as discussed in this paper. The first improvement models signal interaction with multi-faceted surfaces and constructs the return signal at the detector. The second improvement implements several methods for signal detection. These improvements enable the determination of multiple return data and the exploration of radiometric phenomena including dropouts, false alarms, range error, and walk error.

[1] Approved for public release; distribution is unlimited.

2. CONSTRUCTING THE RETURN SIGNAL AT THE DETECTOR

The first step is to construct the return signal at the detector. First, a Gaussian pulse is assumed to be emitted by the transmitter. The pulse is defined by the laser pulse repetition frequency (PRF), average laser power and full-width half-max pulse width (FWHM). The modeled pulse energy distribution across the beam footprint is sampled with a hexagonal pattern to form an *oversample bundle*. A two-dimensional Gaussian shape across the footprint is employed to approximate the power contained within each sample. The power contained by each sample is a fraction of the original emitted Gaussian pulse. The range to, and the reflectance of, a portion of the target is measured individually for each sample. A waveform return is then estimated for each sample. Finally, these sample waveforms are time delayed according to range-to-target then summed together to form a return signal at the detector. This signal is represented in the simulation in terms of optical power because is has yet to go through any of the receiver electronics. Once the return signal at the detector is constructed, the effects of the receiver electronics are modeled by filtering according to the specified receiver components and parameters. The output of the receiver is a voltage waveform that can then be discriminated, according to the desired method, to determine detection points.

2.1. Defining the Emitted Gaussian Pulse

The first step in constructing the return signal waveform at the detector is modeling the emitted pulse from the transmitter. The simulation allows the user to specify its laser pulse repetition frequency (PRF), average laser power, and pulse width. The emitted ladar pulse is modeled as Gaussian and has the same area as the rectangle defined by the FWHM pulse width and pulse power as shown in Figure 1.

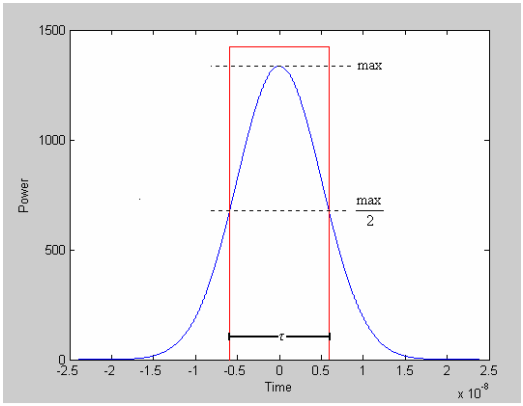

Figure 1 Approximated square pulse with its Gaussian equivalent.

The transmitted power of the Gaussian pulse is defined as

$$P_t(t) = h \cdot e^{\left(-\frac{1}{2}\cdot\left(\frac{t-t_0}{\sigma}\right)^2\right)}. \quad (1)$$

The adjusted values of sigma and the height are given by

$$\sigma = \frac{\tau}{2\cdot\sqrt{2\cdot\ln(2)}}, \quad (2)$$

and

$$h = LPE^* \cdot \left(\frac{2}{\tau}\cdot\sqrt{\frac{\ln(2)}{\pi}}\right), \quad (3)$$

where τ is the pulse width, LPE^* is the laser pulse energy, and t_0 is the pulse delay.

Therefore, the total transmitted power of the Gaussian pulse is

$$P_t(t) = LPE^* \cdot \frac{2}{\tau}\sqrt{\frac{\ln(2)}{\pi}} \cdot e^{\left(-4\cdot\ln(2)\cdot\left(\frac{t-t_0}{\tau}\right)^2\right)} \quad (W). \tag{4}$$

The relationship between the energy and the pulse width reveals that the peak power increases as the pulse width gets shorter. This is because the same amount of energy is essentially being compressed into a shorter pulse. Physics dictate that a direct-detection ladar system has better performance for a given energy pulse as the pulse width becomes shorter.

2.2. Predicting the Returned Pulse

With the emitted temporal pulse defined, a method is introduced that describes how that energy is reflected off multi-faceted surfaces to create a returned temporal pulse. At relatively close distances, ladar beam divergence (β) may be neglected. However, this divergence becomes important in describing the behavior of a system at large distances. The resulting diverged beam of laser energy is modeled by spatially sampling a number of points within the cross section of the beam. As the number of samples increases, so will the spatial accuracy of the signal waveform.

In order to get optimal area coverage within the circular beam cross section, a hexagonal oversampling process was chosen. The relative positions for these hexagonal samples within the beam footprint may be found by setting the area of the circular beam footprint equal to the summed area of the sampled hexagons. This equation is primarily dependant upon the laser beam divergence and the number of desired samples. As the number of samples increase, the vertical (α) and horizontal (γ) angles between them decrease according to:

$$\alpha = \frac{\beta}{2}\cdot\sqrt{\frac{\pi\cdot\sqrt{3}}{6N}} \text{ (rad)}, \tag{5}$$

and

$$\alpha\cdot\sqrt{3} = \gamma \text{ (rad)}, \tag{6}$$

where N is the number of samples.

Figure 2 shows the locations and angle differences of 19 sample points inside a ladar beam cross section using 19 hexagons.

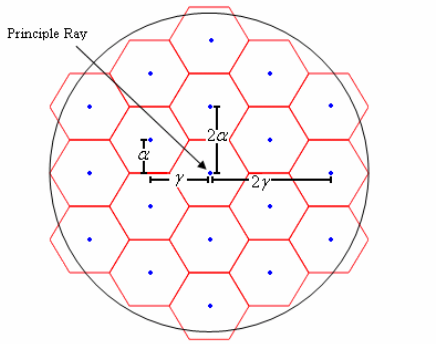

Figure 2 Sampling bundle diagram for 19 samples.

The energy distribution across the cross sectional area of the beam footprint follows a two-dimensional Gaussian shape where the highest concentration of power is contained in the center of the beam and decreases towards the outside edges as seen in Figure 3. The profile is given by

$$P(x,y) = A\exp\left[-2\left(\frac{x^2+y^2}{\beta^2}\right)\right], \tag{7}$$

where A is a scaling constant.

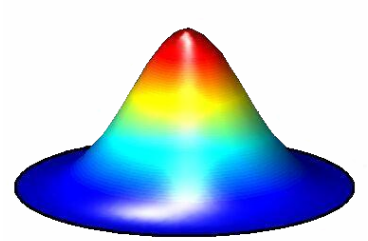

Figure 3. Power distribution across a circular ladar beam cross sectional area.

Given this Gaussian spatial profile, a power *scaling factor* a_{x_i,y_i} is calculated for each hexagonal subsample. The magnitude of the factor is calculated given the number of subsamples and their positions within the beam's cross section as given by

$$a_{x_i,y_i} = \frac{1}{A_T}\exp\left[-2\left(\frac{x_i^{\,2}+y_i^{\,2}}{\beta^2}\right)\right], \tag{8}$$

where

$$A_T = \sum_{i=0}^{N-1}\exp\left[-2\left(\frac{x_i^{\,2}+y_i^{\,2}}{\beta^2}\right)\right], \tag{9}$$

and (x_i, y_i) is the location of the i^{th} sample within the footprint.

This scale factor reflects the amount of power contained in that particular sample of the beam. These weighting factors are then applied to each temporal pulse associated with a sample.

Each sample hits a different part of the target within the beam footprint and is treated as an independent pencil beam. The range to target is known in the simulation and the same emitted Gaussian pulse shape is assumed for each. The returned pulse for each sample arrives at a different time at the receiver, depending on the round-trip distance for each pulse. The magnitude of each returned pulse is scaled, given the power distribution across the emitted pulse footprint.

The round-trip time for the i^{th} sample is calculated by

$$T_i = \frac{2\cdot R_i}{c} \quad (s), \tag{10}$$

where T_i is the time offset for the ith sub-sample, R_i is the range to the object that the ith sub-sample beam strikes, and c is the speed of light.

Using the round trip times, the return pulses of the samples are placed at their appropriate delays on the time line and then summed together at each time step to form the predicted return signal waveform in watts. Figure 4 shows an example of this process for a beam interacting with multiple surfaces.

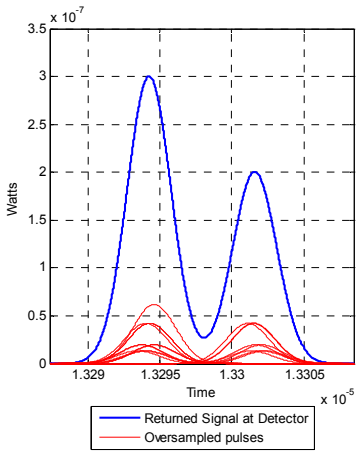

Figure 4. Signal construction from multiple surface returns.

When a ladar beam interacts with an object, the amount of power that is reflected back depends upon two factors, the reflectivity value of that surface and the amount of energy that the material reflects back at a given angle of incidence. These two parameters are described graphically according to Figure 5 [1].

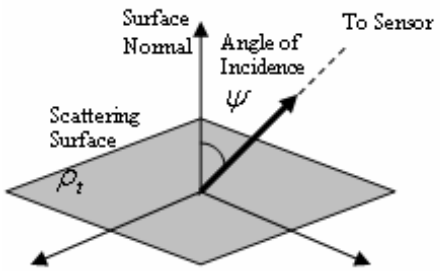

Figure 5. Diagram showing the parameters associated with the bi-directional reflectance distribution function.

The reflectivity of the material ρ_t is dependant solely upon its properties. It is a value between 0 and 1 which represents the amount of energy that is reflected off of the surface to be returned. The returned signal beam width $\Theta_r(\psi)$, represents the area over which the surface radiates the reflected energy and changes according to the angle of incidence. LadarSIM uses these parameters to describe the BRDF factor as

$$BRDF(\psi) \equiv \rho(\psi) = \frac{\rho_t}{\Theta_r(\psi)} \quad (sr^{-1}), \tag{11}$$

where ψ is the angle from incident vector to the surface normal.

LadarSIM has the flexibility to offer the user the option of using a uniform reflectance model or a more complex one. For the uniform reflectance model, the user simply specifies a reflectivity value and chooses whether diffuse, Lambertian, or specular modeling is desired. The reflectivity value entered is treated as a uniform characteristic of all objects in the entire scene. The returned beam width is set to 2π or π when diffuse or Lambertian modeling is selected, respectively. If specular modeling is selected, the Blinn-Phong model is used [2], given by

$$\rho(\psi) = \frac{\rho_t}{\pi} + k_s \cdot \cos^{n-1}\psi\,, \quad \left(sr^{-1}\right), \tag{12}$$

where k_s is the reflectivity at normal incidence, and *n* determines the strength of the specular reflection..

In a real scene, however, there are many different materials and many different angles at which the ladar rays interact with these materials. Each material will have distinct reflectivity values particular to its type and color. The complex reflectance model in LadarSIM allows additional information regarding the reflectivity and the reflectance distribution functions of the included materials to be specified. Each triangle facet in a scene can contain a respective reflectivity value and an index value declaring the type of material it is. These values are then used to determine a much more accurate BRDF factor in describing the amount of energy that is reflected back to the detector.

2.3. Filtering

With the return signal at the detector constructed, as shown in Figure 4, an accurate representation of the optical signal from the reflecting surface has been described. This is the approximate power signal that arrives back at the detector. This signal will then be filtered and processed according to the receiver electronics. A block diagram of the receiver processing is given in Figure 6.

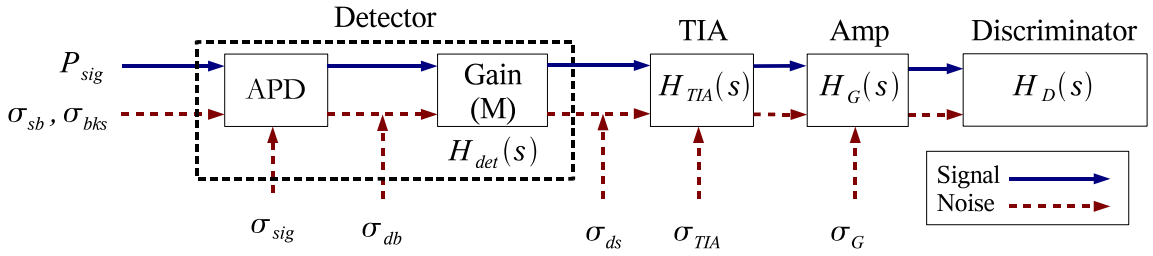

Figure 6. Block diagram showing the transfer functions, signal, and noise components in the receiver.

In this figure, the effects of the photodiode, transimpedance amplifier (TIA), gain amplifier, and timing discriminator are represented. In addition, the additive noises for each processing element are given. The waveform at the discriminator is determined by processing the optical signal on the detector, $P_{sig}(t)$, using the transfer function given by

$$H(s) = H_{\det}(s)H_{TIA}(s)H_G(s)H_D(s), \tag{13}$$

where

$$H_{\det}(s) = \frac{M\psi\omega_{\det}}{s+\omega_{\det}} \equiv \text{ the transfer function of the photodiode,} \tag{14}$$

$$H_{TIA}(s) = \frac{G_{TIA}\omega_{TIA}}{s+\omega_{TIA}} \equiv \text{ the transfer function of the transimpedance amplifier,} \tag{15}$$

$$H_G(s) = \frac{G_G\omega_G}{s+\omega_G} \equiv \text{ the transfer function of the power amplifier,} \tag{16}$$

$$H_D(s) = \frac{\omega_D}{s+\omega_D} \equiv \text{ the transfer function of the discriminator matched filter;} \tag{17}$$

And M is the APD gain of the photodiode, ψ is the detector responsivity at M=1, $\omega_{\det}$ is the bandwidth of the photodiode (rad/s), ω_{TIA} is the bandwidth of the TIA (rad/s), G_{TIA} is the transimpedance gain of the TIA (Ohms or

V/A), ω_G is the bandwidth of the power amplifier (rad/s), G_G is the gain of the gain amplifier (dimensionless), and ω_D is the bandwidth of the discriminator matched filter (rad/s).

LadarSIM allows the user to specify the necessary parameters for the various detector stages such as gains, electrical bandwidths and noise levels. The output of $H(s)$ is the voltage signal that is sent to the discriminator in order to determine the possible detection points. These detection points are then used to determine the signal-to-noise ratio (SNR). This ratio becomes important in predicting whether or not the detection will actually occur at that specific point. This prediction is described most commonly as the probability of detection or dropout.

The range error for each detection point also depends upon the filtered signal. This error is the predicted inaccuracy inherent to the detection process and is dependent upon the slope of the pulse at the point of the detection [3]. As the slope of the pulse increases the predicted range error decreases accordingly. Although these are simply statistical estimates, they are efforts to model the random phenomena seen in actual ladar systems.

3. DISCRIMINATION METHOD IMPLEMENTATION

Introducing waveform processing into LadarSIM not only helps increase the fidelity of the simulation model but it also provides an opportunity to implement and analyze various discrimination methods common to contemporary ladar systems. The discrimination methods available in LadarSIM are a leading edge, constant fraction and a derivative crossover method. Each method carries with it inherent benefits and short-comings that will be discussed.

3.1. Leading Edge Method

The leading-edge discriminator method declares the range value at the time where the return signal crosses a specified voltage threshold. This threshold is proportional to the predicted base noise level of the system that is dependant solely upon the specified detector noises in the simulation. The user can specify the magnitude of this scaling factor in order to achieve the desired performance. In Figure 7, the detection threshold was set for a factor of 5.5 times the noise level, which is indicated by the dashed green line. The point where the returned signal intersects the threshold is indicated by the vertical black line. The range value is determined from this time value.

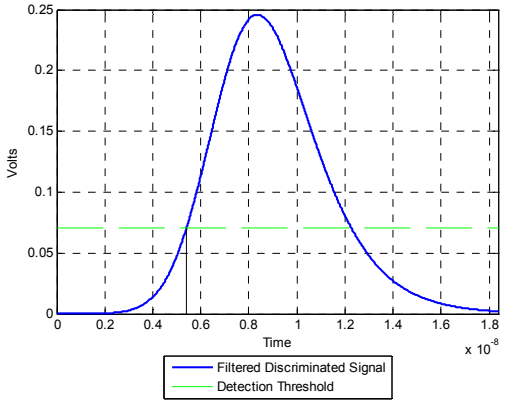

Figure 7 Time determination using the leading edge detection method.

Though the leading-edge discriminator is a simple and effective method to implement, it is susceptible to walk error. Walk error is defined as the error that results from the change in timing due to differences in the return signal shape and amplitude. This can lead to incorrect and misleading range values. For example, if the same pulse energy is reflected off of two surfaces of varying reflectivity, the returned pulse amplitudes will be different. This variation in amplitude will result in the two pulses crossing the threshold at different times, as shown in Figure 8. Thus, the range measurement is biased when high-reflectivity targets are detected. This effect leads to the familiar "floating cross"

effect that is observed in ladar range calibration using highly reflective cross-shaped painted targets on a darker background.

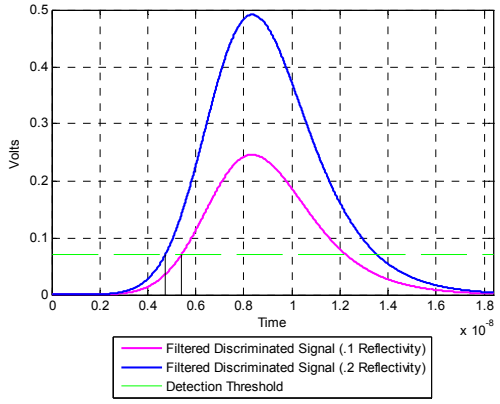

Figure 8 Walk error caused by different pulse amplitudes. The difference in timing corresponds to 0.1019 m of range error.

3.2. Constant Fraction Method

The constant fraction discriminator (CFD) reduces walk error by declaring the range value at the time when a delayed version of the received signal crosses a scaled version of the same signal, as illustrated in Figure 9. Since the intersection point is invariant to pulse amplitude, walk error is essentially eliminated. The crossover point is determined by the delay and fraction (scale) in the discriminator. The delay and fraction are usually chosen to allow the original and delayed signals to cross at the point of highest slope in the delayed signal.

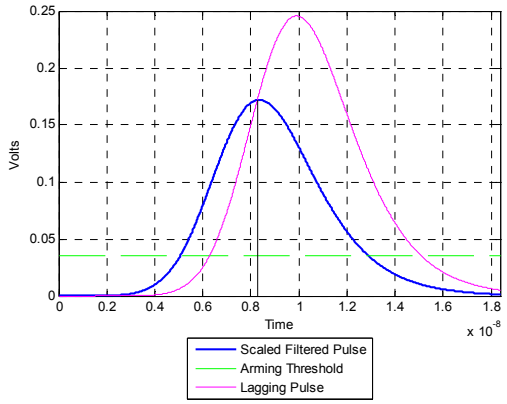

Figure 9 Time determination using the constant fraction method.

An interesting variation of the CFD results when the fraction is set to 1 and the delay is set so that the trailing edge of the pulse crosses the leading edge of the delayed version at the point of the highest slope, as shown in Figure 10. This CFD, known as a lead-lag discriminator, is used because the signal-dependent noise is greatest at the peak of the pulse, and there may be instability in the peak caused by relaxation oscillations in the laser, both of which add to range inaccuracies [4].

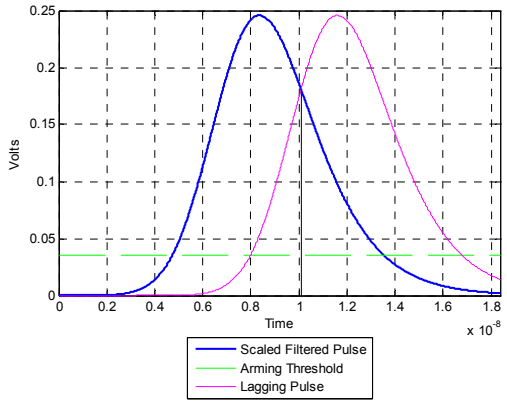

Figure 10 Time determination using the lead/lag method.

In order to reduce false alarms, this discriminator is not enabled until the received signal exceeds a threshold known as the "arming threshold." This threshold is set to result in a desired false alarm rate, but reduces the effective SNR since the crossover must occur while the input signal is above the threshold.

3.3. Crossover (Derivative) Method

The crossover discriminator is another amplitude-invariant discriminator. It has the advantage over the CFD in that it does not need delay lines for implementation and is thus well suited for integration in ASICs [5]. The disadvantage of this method is that is sensitive to walk error due to changes in pulse shape. This method inserts a high pass (HP) filter into the discriminator to act as a differentiator. The HP discriminator declares the range value at the time the HP signal crosses zero. To reduce switching noise, common implementations of this discriminator employ differential HP filters [4]. In this case, the timing point is declared when the differential signals cross, as shown in Figure 11.

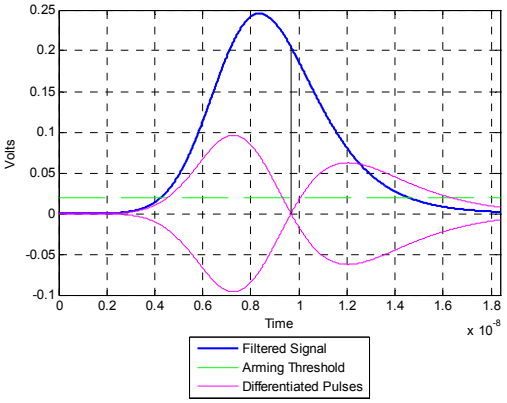

Figure 11 Time determination using the crossover method.

4. MULTIPLE RETURNS

Another benefit of waveform processing is the ability to declare multiple detections. Multiple detections occur normally in actual ladar systems because of diverged ladar beam footprints hitting multi-faceted surfaces causing split energy returns. If these split energy returns contain sufficient energy in each pulse return, multiple detections can be determined. Figure 12 shows an example return signal that could yield two possible detection points when discriminated with the leading edge detection method.

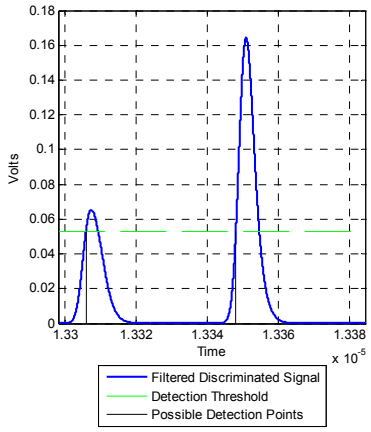

Figure12 Multiple detections determined using the leading edge method.

Data sets containing multiple returns can lead to higher resolution images when viewed in a 3-dimensional point cloud. These images become increasingly more valuable when the desired object being scanned lies under trees, shrubbery or any other concealing cover. In scenes such as these, the first return data will consist mostly of the overlaying cover while the returns which occur after will contain the desired objects.

5. CONCLUSION

With any computer simulation, the tradeoff between accuracy and speed must always be addressed. Since LadarSIM already had and adequate model for ladar beams that hit flat surfaces, this model can still be employed in those areas. According to the desire and time available, the user may decide to utilize waveform analysis for every ladar shot in a scene or have it be focused only in areas of interest. Areas of interest are defined as areas where the simple model is inadequate or around complex objects in the scene. This allows the user the control in finding the balance between accuracy and efficiency.

The implementation of waveform processing within LadarSIM has greatly increased and enhanced the capabilities of the simulation. Using an efficient method, the return signal at the detector is constructed using a process of summing oversampled pulses. This constructed pulse can then be filtered and discriminated according to the specified detector electronics. The detection methods contained within the software package and discussed in this paper all have advantages and disadvantages in terms of statistical detection probabilities and range errors. Having a simulation containing both waveform modeling and various different detection methods becomes a valuable tool in exploring the radiometric performance seen in ladar systems today.

REFERENCES

[1] "IRMA User's Guide, Volume 2" pg. 8-30

[2] J. Blinn, "Models of Light Reflection for Computer Synthesized Pictures," Computer Graphics, Vol. 11, No. 2, July 1977, pp. 192-198

[3] G. Bertolini, and A. Coche, "*Semiconductor Detectors*," North-Holland Publishing Co., Amsterdam, 1968, pp. 243-276

[4] A. Kilpelä, J. Ylitalo, K. Määttä, and J. Kostamovaara, "*Timing Discriminator for Pulsed Time-of-Flight Laser Rangefinding Measurements*," Review of Scientific Instruments, Vol. 69, No. 5, May 1998, pp. 1978-1984

[5] A. Kilpelä, "*Pulsed Time-of-Flight Laser Range Finder Techniques for Fast, High Precision Measurement Applications*," PhD Dissertation, Department of Electrical and Information Engineering, University of Oulu, Finland, 2004.