A Novel Approach to UWB Millimeter High Resolution Range Detection

Robert D. Sims III M.S.
Electrical and Computer Engineering
Member, IEEE
Milford, MI, USA

Daniel N. Aloi Ph.D.
Electrical and Computer Engineering
Senior Member, IEEE
Rochester, MI, USA

Jia Li Ph.D.
Electrical and Computer Engineering
Senior Member, IEEE
Rochester, MI, USA

***Abstract*— The Non-linear Gaussian Recursive Algorithm (NGRA) is a novel algorithm that solves some of the known short range radar impairments. These short range radar impairments include clutter, overlapping echo pulses, antenna pulse distortion, and poor distance resolution. The NGRA accurately models the source signal and subtracts it from the original signal which allows for additional peaks to be detected. In order to model the primary signal and its side lobes, a sum of Gaussian model was chosen. To estimate the models coefficients a non-linear fit algorithm is required using initial conditions generated from a peak detector. The coefficients from the model provide location information and distance resolution beyond the limitations of the sampling rate of the captured data. Through experimentation the NGRA algorithm was proven to be accurate and reliable, achieving between one to eight percent error rates.**

I. INTRODUCTION

Radar algorithms are used for a wide variety of applications. Radar can be used for medical imaging, ground penetrating radar, air traffic control, and a wide range of military applications. Medical imaging uses radar algorithms to process the return pulse and develop imaging of an object and its location. One of the most recent areas of study is ultra-wide band (UWB) breast cancer detection of tumors [4, 14-15]. Traditional x-ray mammography results in 15% of all breast cancer present at the time of screening, and nearly three-fourths of all breast lesions biopsied turn out benign [14]. One method being proposed to solve these issues of traditional x-ray mammography is UWB breast cancer detection. This paper concentrates on solving some of the range estimation issues of medical imaging. In order to understand the issues with radar it is important to understand the echo return pulses from an object, antenna systems with the proper bandwidth, and digital signal processing algorithms to detect pulses and range calculation.

Radar systems use digital signal processing algorithms in order to detect the presence of a pulse in a return signal. In some cases this task is challenging especially if the echo signal amplitude is close to the amplitude of clutter or noise. This typically results in false detections of echo pulses, and advanced signal processing to reduce the false detection rate. In short range radar systems it is possible for echo signals to overlap with themselves and with the source pulse. This is caused by the close proximity of the source and receive antenna being close to the boundaries. Another unique issue to short range radar applications is the sampling frequency limit of the data accusation equipment. For instance a 40 GHz sampling frequency translates to a 7.5 mm distance between samples if the targets are millimeters apart this can cause poor range estimation. In application of medical imaging this can be a critical problem.

This paper proposes a novel range detection algorithm Non-Linear Gaussian Recursive Algorithm (NGRA) that will solve some of these short range radar issues, but can also be used in longer range applications. This is unique compared to algorithms defined in [1-7]. The CLEAN process as described in [7] uses a Fourier transform to find the largest magnitude spot in the signal. After the largest magnitude spot is identified the complex amplitude and position is used to develop a model that is subtracted from the source signal resulting in a new signal devoid of side lobes. NGRA uses a sum of Gaussian model instead of Fourier transform, and can model the side lobes of the radar pulse; CLEAN algorithm requires side lobe reduction. NGRA uses a peak detector to provide peaks of overlapping pulses that the traditional Maximum point (MAX) and interpolation (INT) peek algorithms cannot provide. The Non-homogeneous Poisson process (NHPP) [1] model based algorithms reduce false rejection but when pulses overlap they are less effective. The NGRA has better modeling capability than the NHPP algorithms.

II. NGRA ALGORITHM

The NGRA is an iterative algorithm which models the first pulse and removes it from the source waveform. Through this iterative process the model coefficients are stored and used for range calculation. There are two ways of utilizing this algorithm. The first is using the sum of Gaussian model to model the source pulse signal without targets present and then subtracting it from the signal that contains the boundary reflections. This option is useful when the target completely overlaps with the source signal. The second method models the source signal with the targets present in the signal and then subtracting it to detect additional boundaries. This second method is useful when the echo pulse does not overlap with the source pulse. The NGRA process is defined as follows:

1) Record multiple observations of the source pulse without targets present.
2) Generate initial conditions for non-linear model fit using peak information obtained from the peak detector.

978-1-4799-7088-9/14/$31.00 ©2014 IEEE

3) Perform non-linear fit using initial conditions in step 2.
4) Store model coefficients for future range estimation.
5) repeat steps 2 through 4 for all observations
6) Calculate the mean of the model coefficients stored in step 4
7) Create signal mode from model coefficients in step 6.
8) For each observation containing targets subtract the model created in step 7.
9) repeat steps 2 through 4
10) Calculate Range using model coefficients from step 6 and 9.
11) Repeat steps 6 through 11 until no other targets are present.

For option 2 you can skip steps 1- 5 in the NGRA process. Prior to processing the source signal it is recommended that the observed signals are filtered with a low pass filter. This step will reduce the high frequency noise and remove aliasing that may have been caused by the antennas. The cut off frequency of the filter should be based on the maximum frequency range of your antenna.

A. Peak Detection Algorithm

The peak detector used in the NGRA algorithm is not a traditional peak detection method. This method allows both peaks and valleys of the signal to be detected providing more information to be used as initial conditions. This information provides improved the convergence time and accuracy of the non-linear fit. The first step is to calculate the derivative of the source signal using equation 3.2. V is the derivative of the signal S. Where n is the sample vector and N is the total number of samples in the sampled signal.

$$V(n) = \frac{s(n+1)-s(n)}{\Delta n}, n = 1,2,3 \dots (N-1) \qquad (3.2)$$

Δn is the difference between sample values which is equal to 1. Points of inflection can be calculated from V(n) by detecting the change in sign of adjacent V(n) values. In order to prevent false detection of clutter and noise in the signal two thresholds must be applied. The first threshold limits the peak detection by amplitude of the peak, T1. This threshold will help to remove the noise and clutter from the signal. The second threshold T2 is used to remove noise that may be riding on the signal. T2 is the minimum distance between two peaks that should be detected. It is also desirable that there is a zero crossing between these two peaks. These thresholds are defined by a series of detection equations that are defined in this section. Unless otherwise specified equation 3.3 defines n.

$$n = 1,2,3 \dots N-1 \qquad (3.3)$$

The first step in the calculation is to determine inflection points in the signal. Using equation 3.2 to calculate the derivative of the signal the inflection points can be found where two consecutive sample points change from positive to negative or negative to positive. Equation 3.4 generates a detection vector $I_p(n)$ which contain logical values at the points prior to inflection.

$$I_p(n) = \begin{cases} 1, & (v(n) \geq 0 \; and \; v(n+1) \leq 0) or \\ & (v(n) \leq 0 \; and \; v(n+1) \geq 0 \\ 0, & Otherwise \end{cases} \qquad (3.4)$$

After the points of inflection are calculated it is necessary to apply the general amplitude threshold to perform the course filtering for peaks and valleys of interest. Equation 3.5 calculates the detection vector $a(n)$, which contains logical value one for the location of peaks greater than the threshold, and changes sign from negative to positive.

$$a(n) = \begin{cases} 1, & |s(n+1)| > T1 \; and \; I_p(n) = 1 \\ 0, & otherwise \end{cases} \qquad (3.5)$$

Once the first threshold is applied it is necessary to generate a location vector of the peaks prior to removing the zero spacing in the amplitude vector. Equation 3.6 generates the first location vector

$$l_p(n) = \begin{cases} n, & a(n+1) = 1 \\ 0, & otherwise \end{cases} \qquad (3.6)$$

In order to remove the zero value in the source vector a diagonal matrix is first generated from the values in equation 3.5 as shown in equation 3.7.

$$I_d = diag(a) \qquad (3.7)$$

After the location vector diagonal matrix is generated it is possible to perform matrix multiplication of the location vector against the signal vector to obtain the magnitude, equation 3.8. It is also necessary to capture the reorganize the location vector in equation 3.9.

$$p_v(n) = I_d * s \qquad (3.8)$$

$$p_l(n) = I_d * l_p \qquad (3.9)$$

This calculation reorganizes the data in 3.8 and 3.9 to have all the locations in the beginning of the vector. This is accomplished with using the diagonal matrix created by the location vectors. By having ones or zeros across the diagonal matrix, it effectively selects the detected peaks data and moves it to the beginning of the vector. This is necessary preparation for the zero crossing detection and the amplitude difference calculation. The next step is to calculate the zero crossing detector vector using equation 3.10. Zero crossing determines if two adjacent peaks have different signs. If they are different in sign it is considered to cross zero.

$$z(n) = \begin{cases} 1, & (p_v(n) > 0 \; and \; p_v(n+1) \leq 0) \\ & or \; (p_v(n) < 0 \\ & and \; p_v(n+1) \geq 0) \\ 0, & Otherwize \end{cases} \qquad (3.10)$$

The next stage of the calculation is to detect adjacent peaks that are not far enough apart in magnitude, equation 3.11. When performing this level of detection it is necessary to take into account the zero crossing detection in the calculation.

$$d(n) = \begin{cases} 1, & |p_v(n) - p_v(n+1)| > T2 \text{ or } z(n+1) = 1 \\ 0, & Otherwize \end{cases} \quad (3.11)$$

Given that the previous detection step removed peak locations, it is necessary to reorder the detection vector again. This will be done first by creating a diagonal matrix using the detection vector d, equation 3.12

$$I_d = diag(d) \quad (3.12)$$

After the diagonal matrix I_d has been generated it is possible to use matrix multiplication again to select the detected peak data as calculated in 3.13, and 3.14.

$$p_{v2} = I_d * p_v \quad (3.13)$$

$$p_{l2} = I_d * p_l \quad (3.14)$$

Since the data still contains trailing zeros from the calculations performed in 3.13 and 3.14 it is good to remove them from the vector. Since the detection vector is comprised of zeros and ones it is possible to calculate its length by summing the vector. The sum of the vector is calculated using equation 3.15.

$$N_p = \sum_{n=1}^{N-1} d(n) \quad (3.15)$$

After the number of peaks is calculated it is possible to remove the trailing zeros by the calculations in equation 3.16 and 3.17.

$$P = p_{v2}(n),\ n = 1,2,3 \ldots N_p \quad (3.16)$$

$$L = p_{l2}(n),\ n = 1,2,3 \ldots N_p \quad (3.17)$$

The information from equations 3.16 and 3.17 will be used to provide initial conditions to fit the pulse model.

B. Pulse Model

In order to provide better resolution and to suppress the effects of side lobes in the signal, the sum of Gaussian model has been chosen. This model is different than the one used in NNHP. The NNHP models used a sum of a constant and a Gaussian. The new model proposed in this paper is a sum of Gaussians without a constant. Using a sum of Gaussian model allows for the modeling of the side lobes of the signal which provided the inspiration for this model. This model accurately models the distorted pulse resulting from transmission though an antenna. The Gaussian function is defined in equation 3.18.

$$f(n) = a_0 e^{-\left(\frac{n_0-\mu_0}{u_0}\right)^2} + a_1 e^{-\left(\frac{n_1-\mu_1}{u_1}\right)^2} \ldots a_k e^{-\left(\frac{n_k-\mu_k}{u_k}\right)^2} \quad (3.18)$$

where n is the sample number a, μ, and u are the coefficient's being fitted and k=0,1,..K K is the total number of Gaussians in the model. A sum of Gaussian model is a non-linear function and requires a non-linear fitting method. The method used to fit this model is a non-linear least mean squares trust-region-reflective [9-13], using MATLAB. The equation for this type of fit is defined in equation 3.19.

$$\min_x \|F(x, \hat{y}) - y\|_2^2 = \min_x \sum_i (F(x_i, \hat{y}_i) + y_i)^2 \quad (3.19)$$

Input data from the measurement is vector $\hat{y}$, and y is observed output data vector of the fit function. The goal of this equation is to minimize the error between the predicted data and the measured data by adjusting the values in coefficient vector x. When using non-linear fitting methods, it is important to choose initial conditions that are as close as possible to the target values in order to reduce the number of step's that the non-linear fitting algorithm takes to converge on the solution. It is also important to focus the fit on the section of the signal you are trying to fit. If you pick initial conditions far away from the desired pulse, it will fit a section that is not desired. In order to start with good initial conditions data from the peak detector will be used. The locations of the peaks will be used as the mean coefficient initial condition and the amplitude at the peak locations will be used for the corresponding amplitude coefficient. The peak detector needs to capture enough peaks to provide initial conditions for the model. This is accomplished by adjusting the T1 and T2 thresholds appropriately. Initial condition for the variance is considered to be a value of n=1.

C. Range Calculation

Range is calculated using model coefficients generated by fitting the model defined in equation 3.29. The μ corresponding to the primary peak is used to calculate the range between peaks. It is important to use the correct coefficient value that represents the primary peak and not the side lobe of the peak. This peak corresponds the mean of the second Gaussian in the sum of 3 Gaussian model. Using the fitted coefficient instead of the sample value from the peak detector will provide a more accurate range resolution than can be offered by the minimum resolution of the sampling frequency. The following equation is used to calculate the range of the boundaries:

$$R(m) = \left(\hat{X}_m(5) - \hat{X}_0(5) - \frac{\sum_0^M R(m)}{\frac{c}{fs*\sqrt{\mu_r \varepsilon_r}}}\right) * \frac{\frac{c}{fs*\sqrt{\mu_r \varepsilon_r}} d}{2} \quad (3.29)$$

Where c is the speed of light, μ_r is the relative permittivity, and ε_r is the relative permeability of the materials the wave travels through. m=1,2,…M where M is the total number of coefficient's in the $\hat{X}$. $\hat{X} = [a_0, \mu_0, u_0, a_1, \mu_1, u_1, \ldots M]$. M is dependent on the number of fitted peaks in the signal.

III. Experimental Setup

The materials used in this experiment are analogous to those found in the human body. Soybean oil is used to mimic the fatty tissue in the body, and the acrylic cube was chosen to reduce the reflection between the oil and the container [14]. The aluminum-foil and copper represent materials that could be shrapnel from an explosion in the body or part of an implant. The latex balloon filled with ketchup is a simple homogeneous model analogous to a malignant tumor, having similar dielectric constants as described in [15].

In order to verify accuracy of the NGRA algorithm a test setup analogous to the human body was chosen. An acrylic cube containing a homogenous material, soy bean oil, to mimic

human fatty tissue. Targets were chosen to provide different levels of frequency. Aluminum foil, Copper FR4 cladded board, and a latex balloon are the boundaries chosen for the experiment. This provided echo pulses with different levels of echo signals. Copper of dimensions 115 mm x 150 mm provided the strongest echo signals. . Next, an Aluminum bar of dimensions 22mm x 13 mm x 270 mm provided a lower amplitude echo pulse do to its small cross section. A mid-level amplitude pulse generated by the balloon filled with ketchup with a diameter of 40 mm. Since this experiment is focused on algorithm development the skin boundary was excluded. These targets provide a range of permittivity and will produce reflections of different magnitude. The test setup illustrated in Figure 1 was used to generate and capture reflections of a UWB pulse from a target submerged in soybean oil. The horn antenna was chosen for the source antenna due to its large gain and bandwidth. The SAS-571 from AH Systems has a bandwidth range of 700MHz and 18 GHz. For the receive antenna a checkerboard patch antenna was created from FR 4 board based on the optimal design found in [8]. This antenna provides a 3.1GHz to 10.6 GHz band width also providing a small footprint avoiding interfering with the source pulse. In order to develop a wide band pulse the picosecond pulse generator with a FWHM of 70 ps Gaussian pulse at -6.72 V. The pulse is generated negative in voltage, but will be positive when transmitted through the antenna.

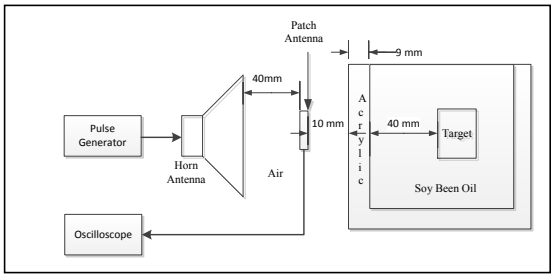

Figure 1. Experimental setup diagram.

IV. Results

This section shows the results of original pulse modeling, peak detection, and estimation of range using the NGRA algorithm. The first evaluation was performed by modeling the source pulse and subtracting it from the collected reflections. Peak detection of the first and secondary boundaries is shown in Figure 2. And the results of the range calculations are shown in Table 1.

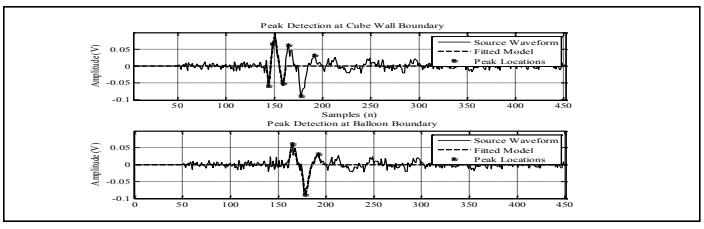

Figure 2. Peak detection of first and second boundary.

The second test performed shows the performance of option number 2 of the NGRA algorithm. Option 2 models the source signal with the targets present in the signal and then subtracting it to detect additional boundaries. Figure 3 shows the fitting of the initial pulse and the range detection of the secondary boundary.

TABLE I. RANGE CALCULATION WITH SOURCE MODELING.

Material	Results			
	Range To Cube(mm)	Range To Cube % Error	Range 2 (mm) Expected 40 mm	% Error Range 2
Aluminum Foil	9.18	8	39.5	1
Copper RF4	9.79	2	40.4	-1
Balloon Latex	10	0	41.1	-3

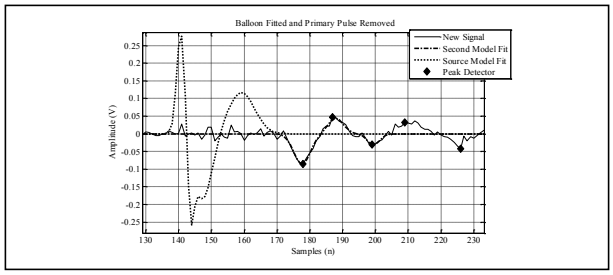

Figure 3. Fitting of initial pulse of 70 mm range estimate.

TABLE II. RANGE CALCULATION RESULTS 70MM TARGET.

Material	Results	
	Range (mm)	*% Error Range*
Aluminum Foil	74.3	-6.1
Copper RF4	70.4	0.6
Balloon Latex	69.1	1.2

V. Conclusion

The NGRA algorithm is effective at improving range estimation of boundary's that are in the millimeter range. This algorithm's unique capability of using peak detection and sum of Gaussian model allow it to overcome short range radar challenges. The application of this algorithm is not limited to the applications investigated in this paper. This algorithm can be used for long range radar with weak radar cross section targets, or even traditional radar applications. The NAGRA algorithm can be applied to real time radar systems, with the appropriate analysis window size. This paper has proven the accuracy of the NGRA algorithm; with a few minor improvements it can be even better. The algorithm has a few short comings as well. One of the short comings of the option 1 implementation is the positioning of the source model, and its sensitivity to test setup changes. These issues can be overcome with better antenna mounts in the future experiments.

REFERENCES

[1] Yıldırım, A.; , "Experimental analysis of various estimators for target range estimation in high resolution radars," Signal Processing, IET , vol.5, no.7, pp.680-689, Oct. 2011.

[2] Yildirim, A.; Efe, M.; Ozdemir, A.K.; , "An Alternative Model for Target Position Estimation in Radar Processors," Signal Processing Letters, IEEE , vol.14, no.8, pp.549-552, Aug. 2007.

[3] Mir, H.S.; Carlson, B.D.; , "On the Definition of Radar Range Resolution for Targets of Greatly Differing RCS," Instrumentation and Measurement, IEEE Transactions on , vol.61, no.3, pp.655-663, March 2012.

[4] Williams, T.C.; Fear, E.C.; Westwick, D.T.; , "Tissue sensing adaptive Radar for breast cancer detection-investigations of an improved skin-sensing method," Microwave Theory and Techniques, IEEE Transactions on , vol.54, no.4, pp.

[5] Wong, A.M.H.; Eleftheriades, G.V.; , "Superoscillatory Radar Imaging: Improving Radar Range Resolution Beyond Fundamental Bandwidth Limitations," Microwave and Wireless Components Letters, IEEE , vol.22, no.3, pp.147-149, March 2012.

[6] Kulpa, K.; , "The CLEAN type algorithms for radar signal processing," Microwaves, Radar and Remote Sensing Symposium, 2008. MRRS 2008 , vol., no., pp.152-157, 22-24 Sept. 2008

[7] Tsao, J.; Steinberg, B.D.; , "Reduction of sidelobe and speckle artifacts in microwave imaging: the CLEAN technique," Antennas and Propagation, IEEE Transactions on , vol.36, no.4, pp.543-556, Apr 1988

[8] Moosazadeh, M.; Ghobadi, C.; Dousti, M.; , "Small Monopole Antenna With Checkered-Shaped Patch for UWB Application," Antennas and Wireless Propagation Letters, IEEE , vol.9, no., pp.1014-1017, 2010

[9] Coleman, T.F. and Y. Li, "An Interior, Trust Region Approach for Nonlinear Minimization Subject to Bounds," SIAM Journal on Optimization, Vol. 6, pp. 418-445, 1996.

[10] Dennis, J. E. Jr., "Nonlinear Least-Squares," State of the Art in Numerical Analysis, ed. D. Jacobs, Academic Press, pp. 269-312, 1977.

[11] Levenberg, K., "A Method for the Solution of Certain Problems in Least-Squares," Quarterly Applied Math. 2, pp. 164-168, 1944.

[12] Marquardt, D., "An Algorithm for Least-Squares Estimation of Nonlinear Parameters," SIAM Journal Applied Math., Vol. 11, pp. 431-441, 1963.

[13] More, J. J., "The Levenberg-Marquardt Algorithm: Implementation and Theory," Numerical Analysis, ed. G. A. Watson, Lecture Notes in Mathematics 630, Springer Verlag, pp. 105-116, 1977.

[14] Xu Li; Davis, S.K.; Hagness, S.C.; van der Weide, D.W.; Van Veen, B.D.; , "Microwave imaging via space-time beamforming: experimental investigation of tumor detection in multilayer breast phantoms," Microwave Theory and Techniques, IEEE Transactions on , vol.52, no.8, pp. 1856- 1865, Aug. 2004.

[15] Alshehri, S.; Jantan, A.; Abdullah, R.S.A.R.; Mahmud, R.; Khatun, S.; Awang, Z.; , "A UWB imaging system to detect early breast cancer in heterogeneous breast phantom," Electrical, Control and Computer Engineering (INECCE), 2011 International Conference on , vol., no., pp.238-242, 21-22 June 2011.

[16] Yildirim, A.; Efe, M.; Ozdemir, A.K.; , "Peak picking losses in radar detectors," Signal Processing and Information Technology, 2005. Proceedings of the Fifth IEEE International Symposium on , vol., no., pp.136-140, 21-21 Dec. 2005.