REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
01/12/2016	FINAL	14/05/2003 - 01/12/2016

4. TITLE AND SUBTITLE
SYNTHETIC APERTURE SONAR PROCESSING WITH MMSE ESTIMATION OF TARGET SAMPLE VALUES

5a. CONTRACT NUMBER
N00014-03-C-0226

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
ALTES, RICHARD, A.

5d. PROJECT NUMBER

5e. TASK NUMBER
A002

5f. WORK UNIT NUMBER
0004

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
CHIRP CORPORATION
8248 SUGARMAN DRIVE
LA JOLLA, CA 92037

8. PERFORMING ORGANIZATION REPORT NUMBER
0226-0004

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
OFFICE OF NAVAL RESEARCH
BALLSTON TOWER ONE
800 NORTH QUINCY STREET
ARLINGTON, VA 22217-5660

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
A002

12. DISTRIBUTION/AVAILABILITY STATEMENT
UNCLASSIFIED/UNLIMITED

13. SUPPLEMENTARY NOTES

14. ABSTRACT
MMSE (minimum mean-square error) target sample estimation using non-orthogonal basis functions is applied to object imaging with SAS/multi-aspect data. Theoretical and simulation results are impressive: Accurate images from echoes at three aspects, without a significant increase in SNR. The MMSE process creates images with new features that can be applied to image-based compensation of receiver rotation/location measurement errors. The simulation, however, indicates that a weighted sum of current image features is insufficient to reliably compensate for such errors when the gradient of the weighted sum with respect to location/rotation parameters is used in an optimization procedure.

15. SUBJECT TERMS
SYNTHETIC APERTURE SONAR, MULTI-ASPECT IMAGING, IMAGE-BASED MOTION COMPENSATION, MULTI-STATIC SONAR, MULTISTATIC RADAR, SONAR TARGET IMAGING, RADAR TARGET IMAGING, SPOTLIGHT SYNTHETIC APERTURE SONAR/RADAR, ANIMAL SONAR

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			RICHARD ALTES
Unclassified/ Unlimited	Unclassified/ Unlimited	Unclassified /Unlimited	Unlimited	22	**19b. TELEPHONE NUMBER** *(Include area code)* (858) 453-4406

Final Report for Contract N00014-03-C-0226

Title: Synthetic Aperture Sonar Processing with MMSE Estimation of Image Sample Values

Author: Richard A. Altes, Ph.D.

Contractor: Chirp Corporation, 8248 Sugarman Drive, La Jolla, CA 92037

Abstract

An innovative technique for efficient SAS and multi-aspect sonar imaging has been developed and tested with a software simulation. A reflecting object is modeled with an array of unit-amplitude point targets with locations corresponding to sample points on a 2-D or 3-D sonar image. Each point target has a predicted multi-aspect echo consisting of delayed signal samples at the elements of a receiving array on the SAS vehicle, at a sequence of vehicle transmit/receive locations. Each multi-aspect echo is used as a basis function, and each point on the object has its own basis function. Although these basis functions usually are not orthogonal, they can still be used in a minimum mean-square error (MMSE) estimator that models the object echo as a weighted sum of the multi-aspect basis functions. The complex-valued weights are estimated and used as image samples. Part of the MMSE estimator forms a conventional (maximum likelihood) SAS or multi-aspect sonar image via space-time matched filtering. The MMSE estimator, however, also includes an image sharpening operation that compensates for basis function non-orthogonality and significantly improves the conventional image, especially for a small number of observations (e.g., three transmit/receive locations).

The processor is vulnerable to vehicle location/rotation measurement errors that degrade the images. Various features from maximum likelihood and MMSE images have been tested for image-based error correction. Although some promising features have been identified, their gradients with respect to various measurement errors are not sufficiently reliable for consistent location/rotation error compensation. The current estimator can be used without error compensation for one-pulse multi-static imaging with one fixed transmitter/receiver and two more fixed receiving arrays. This configuration can image underwater intruders with a single transmission and can be applied to multi-static radar imaging. The estimator may be a viable model for animal sonar, especially if the animals have learned to solve the location/rotation error correction problem.

Introduction

Minimum mean-square error (MMSE) estimation is applied to target imaging with synthetic aperture sonar (SAS). The estimated parameters are sample values of target reflectivity. A display of these samples is a sonar image of an object.

Estimated MMSE parameters typically are weights applied to data-independent orthonormal basis functions. The approach given here, however, defines the k^{th} basis function as the unit-amplitude point-target echo from the k^{th} sample location of the imaged object, as observed at all available receiver array elements and aspects. The resulting basis functions usually are not orthogonal. With this basis, part of the MMSE target sample estimator uses the same space-time matched filtering as in standard array processing, including conventional SAS. The full MMSE estimator, however, significantly improves the imaging performance of the standard (maximum likelihood) process, especially for a restricted number of observations from different aspects.

SAS processing synthesizes a large array by coherently processing echoes received by a moving vehicle [1]. SAS mitigates a resolution discrepancy that affects sonar images. Range resolution can be comparatively accurate when wideband signals are used, but azimuth/elevation resolution is degraded by beam spreading when the target range is much larger than the size of the receiving array. Azimuth resolution in radians is proportional to the wavelength at the signal center frequency divided by the aperture (array) length [2], or the reciprocal of the number of wavelengths contained in the aperture. A large synthetic aperture reduces the range/azimuth resolution discrepancy.

SAS can be interpreted as a process that compensates for the range/azimuth resolution disparity by viewing a target from multiple aspects. Good range resolution at one aspect compensates for poor azimuth resolution at a different aspect. Tomographic SAS estimates target reflectivity samples by combining echoes from multi-aspect spotlights that have good range resolution and poor azimuth resolution. The process typically uses the Radon transform [3], which (like conventional SAS) requires data from many different aspects.

Application of MMSE estimation to SAS yields accurate images with echo data from only a few aspects. Utilization of three aspects, for example, is demonstrated in this report. This capability has high payoff for (1) efficient SAS processing, (2) three-dimensional imaging, (3) target imaging with any vehicle trajectory that allows echo reception from a few different aspects (e.g., approaching a target along a zig-zag path or hovering), (4) multi-static radar/sonar imaging, and (5) analysis of animal sonar capabilities.

The MMSE estimator requires highly accurate knowledge and/or tracking of target/vehicle locations and the computational ability to invert high-dimensional complex-valued Hermitian matrices. The multi-aspect images themselves ideally can be used for correction of location/rotation errors via sharpness measures and other features. Progress has been made on image-based error correction, but more research is necessary. This report motivates such work by illustrating the high payoff for MMSE multi-aspect imaging.

MMSE coefficient estimation in a data model consisting of a sum of weighted basis functions

The relevant MMSE estimator pertains to estimation of the coefficients $\{a_k\}$ in a data model consisting of a sum of weighted basis functions:

$$data(t_i) = \sum_{k=1}^{K} a_k b_k(t_i) \quad i=1,\ldots,I_{max};\ \ k=1,\ldots,K \tag{1}$$

where the basis functions $\{b_k(t)\}$ are not necessarily orthogonal.

The mean square error is

$$MSE(a_1,a_2,\ldots,a_K) = \sum_{i=1}^{I_{max}} |data(t_i) - \sum_{k=1}^{K} a_k b_k(t_i)|^2 \ . \tag{2}$$

A necessary condition for minimizing MSE is that the derivative of MSE with respect to a_j equals zero for $j=1,..,K$. This condition is satisfied if

$$\sum_{i=1}^{I_{max}} [data(t_i) - \sum_{k=1}^{K} a_k b_k(t_i)]\,[b_j(t_i)]^* = 0, \quad j=1,..,K \tag{3}$$

or

$$\sum_{i=1}^{I_{max}} data(t_i)[b_j(t_i)]^* = \sum_{k=1}^{K} a_k \sum_{i=1}^{I_{max}} b_k(t_i)\,[b_j(t_i)]^* \quad j=1,..,K. \tag{4}$$

Letting

$$r_j = \sum_{i=1}^{I_{max}} data(t_i)[b_j(t_i)]^* \tag{5}$$

and

$$c_{jk} = \sum_{i=1}^{I_{max}} b_k(t_i)\,[b_j(t_i)]^* \tag{6}$$

(4) can be written

$$r_j = \sum_{k=1}^{K} c_{jk} a_k \ , \quad j=1,..,K\ . \tag{7}$$

Eq (7) corresponds to the matrix equation

$$\boldsymbol{r} = \boldsymbol{C}\boldsymbol{a} \tag{8}$$

where the desired parameters $a_1,a_2,\ldots,a_K$ are the elements of the column vector $\boldsymbol{a}$ and (6) indicates that $c_{kj} = [c_{jk}]^*$ which means that $\boldsymbol{C}$ is Hermitian. The MMSE estimate of the coefficients $\{a_k\}$ is

$$\hat{\boldsymbol{a}} = \boldsymbol{C}^{-1}\boldsymbol{r}\ . \tag{9}$$

The mean-square error associated with this estimate is obtained by adding noise samples to the data vector. Let

$$\boldsymbol{r} = \boldsymbol{Ca} + \boldsymbol{n} \tag{10}$$

where $\boldsymbol{n}$ is a column vector of K independent zero-mean complex noise samples with variance $\sigma_n^2/2$.

Letting $E\{\cdot\}$ denote statistical expected value,

$$\begin{aligned}
MSE &= (1/K)E\{(\hat{\boldsymbol{a}} - \boldsymbol{a})^T(\hat{\boldsymbol{a}} - \boldsymbol{a})^*\} \\
&= (1/K)E\{(\boldsymbol{C}^{-1}\boldsymbol{r} - \boldsymbol{a})^T(\boldsymbol{C}^{-1}\boldsymbol{r} - \boldsymbol{a})^*\} \\
&= (1/K)E\{[\boldsymbol{C}^{-1}(\boldsymbol{Ca} + \boldsymbol{n}) - \boldsymbol{a})]^T[\boldsymbol{C}^{-1}(\boldsymbol{Ca} + \boldsymbol{n}) - \boldsymbol{a}]^*\} \\
&= (1/K)E\{[\boldsymbol{C}^{-1}\boldsymbol{n}]^T[\boldsymbol{C}^{-1}\boldsymbol{n}]^*\}.
\end{aligned} \tag{11}$$

Further insight is obtained by letting K=2, i.e., by estimating the coefficients of two basis functions that may not be orthogonal. If both basis functions have unit energy and their correlation coefficient in (6) is ρ, then the basis function correlation matrix is

$$\boldsymbol{C} = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

In this case [4]

$$\boldsymbol{C}^{-1} = [1/(1-\rho^2)] \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix} \tag{12}$$

and

$$\begin{aligned}
\boldsymbol{C}^{-1}\boldsymbol{n} &= [1/(1-\rho^2)] \begin{bmatrix} 1 & -\rho \\ -\rho & 1 \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} \\
&= [1/(1-\rho^2)] \begin{bmatrix} n_1 - \rho n_2 \\ -\rho n_1 + n_2 \end{bmatrix}.
\end{aligned} \tag{13}$$

The assumed independence of the noise samples n_1 and n_2 implies that

$$E\{n_1 n_2^*\} = 0 \tag{14}$$

and the expected noise power is

$$E\{|n_1|^2\} = E\{|n_2|^2\} = \sigma_n^2/2. \tag{15}$$

Substituting these definitions into (11-13),

$$\text{MSE} = (1/K)[\boldsymbol{C}^{-1}\boldsymbol{n}]^{T}[\boldsymbol{C}^{-1}\boldsymbol{n}]^{*} = \sigma_n^2\,[(1+|\rho|^2)/|1-\rho^2|^2]\,. \qquad (16)$$

The mean-square error increases with $|\rho|$. MSE=σ_n^2 at $|\rho|=0$ (orthogonal basis functions) to infinity at $|\rho|=1$ (perfectly correlated basis functions). This increase is very gradual at small $|\rho|$ values. At $|\rho|=1/5$, for example, MSE=1.13 σ_n^2 and at $|\rho|=1/3$, MSE=1.4 σ_n^2.

Space-time echoes from target points as basis functions

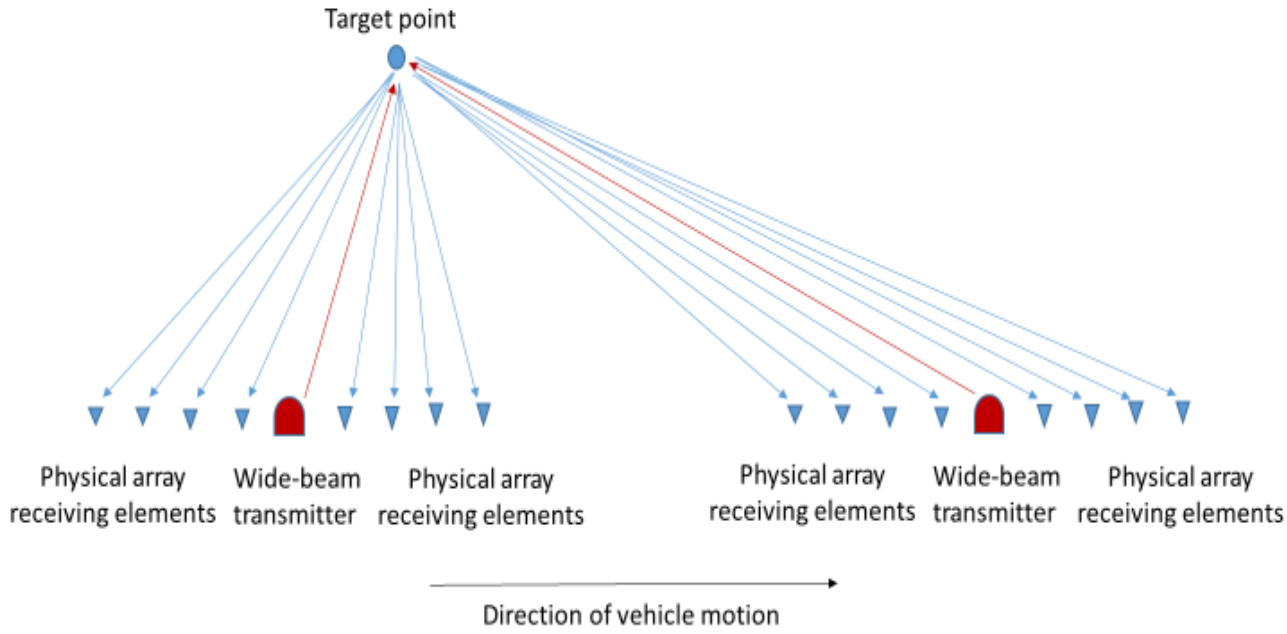
The target/environment model is an array of point-targets or sample points, where the distance between samples is the desired resolution of the sonar imaging system. Although this distance can be different for range and azimuth/elevation, this report uses 2.5 cm (one inch) sample spacing in range and azimuth in a 2-D target model. The unknown parameter $\boldsymbol{\alpha}_k$ is a complex coefficient (target sample value) that multiplies a basis function corresponding to the k^{th} target sample. This basis function is the expected space-time echo from a unit-amplitude point-target at the location of the target sample.

Each point-target echo is measured with a signal transmitted from a small, wide-beam source on the vehicle and received by the physical elements in a relatively small array mounted on the vehicle. The signal is assumed to be frequency modulated or coded so as to increase bandwidth, thus improving range resolution and SNR after the data-basis correlation in (5). The basis functions in (1) include the outputs from all the elements in the receiving array, and are thus functions of array element index (denoting relative element location) as well as time samples of the delayed transmitted signal observed at the output of each element. The correlation of data with basis functions to form the ***r***-vector components in (5) is equivalent to a combination of delay-and-sum beam forming to focus on a target point (spatial matched filtering) and temporal matched filtering with the expected echo time series at each array element (signal samples with delays predicted from the locations of the emitter, the target sample, and the array element).

Although 2.5 cm range resolution can be implemented, the physical array is much too small to achieve 2.5 cm azimuth resolution. Target samples seen from a single aspect are unresolvable if they are within the same range resolution cell and within the array beam pattern, which has width proportional to (range)×(wavelength)/(array length). The resulting ***C***-matrix contains off-diagonal elements that are nearly as large as the diagonal elements, and is ill-conditioned [5]. The desired inverse of this matrix may not exist or may be unreliable. Even if a reliable inverse can be found, $|\rho|$ is close to one and the MMSE estimation error in (16) is large.

The azimuth resolution problem can be mitigated by extending the basis functions to incorporate predicted echoes from multiple aspects. When the vehicle moves to a new aspect as in Fig. 1, a new transmission from the new location results in a new set of space-time echoes from the same target points as before. The new and old basis functions can be combined by adjoining them.

Fig. 1. Construction of a multi-aspect space-time basis function from predicted delays of separate transmitted signals that are received at vehicle array elements at two path positions



The adjoining process is as follows: (1) The old basis function for target point *k* is represented as a vector via a raster scan of all the relevant predicted time samples (corresponding to delayed signal samples) received at array element #1 followed by all relevant predicted, delayed time samples received at array element #2, etc. (2) Similarly, the new basis function for target point *k* is represented by a raster-scan vector with new predicted signal delays corresponding to the new vehicle position. (3) The old and new vectors are adjoined to form an extended basis function vector. (4) The same procedure is applied to observed data, i.e., the observed space-time echo data samples from the old and new vehicle positions are raster scanned and adjoined to form an extended data vector.

The resulting multi-aspect basis functions contain many delayed signal samples, but each such basis function refers to a particular target point. The number of multi-aspect basis functions equals the original number of target points *K*, and the dimension of the ***C***-matrix (*K*×*K*) is unchanged.

The point spread function

It is useful to visualize the estimator response to an isolated, unit-amplitude point-target before applying the $\boldsymbol{C}^{-1}$ sharpening operation. The resulting image represents the $\boldsymbol{r}$-vector in (8) for the one-point target distribution. This image corresponds to the point spread function of the sonar imaging system without sharpening, i.e., to the usual maximum likelihood image that would be formed via coherent space-time matched filtering of echoes from several different aspects. The conventional SAS image is such a maximum likelihood representation, obtained via space-time matched filtering of stored echoes from many observation points along the vehicle path.

The effect on the point spread function of using N observation points is shown in Figs. 2a-2c, where the physical vehicle array is assumed to have a Gaussian beam pattern. Each unit-amplitude Gaussian beam pattern is rotated and added to the others, resulting in a point spread function with a central sharp peak with amplitude N. The peak-to-sidelobe level of the point spread function equals $1/N$, where N is the number of well-separated transmit/receive points along the vehicle path.

A top view of each point spread function resembles an asterisk with a dark dot at its center. Inclusion of more aspects increases the number of line segments in the asterisk and darkens the center point. The smearing effect of the line segments on an image is eliminated by the $\boldsymbol{C}^{-1}$ operation in MMSE estimation of target image samples or (less efficiently) by increasing the number of observed aspects as in conventional SAS and as illustrated in Figs. 2a-2c.

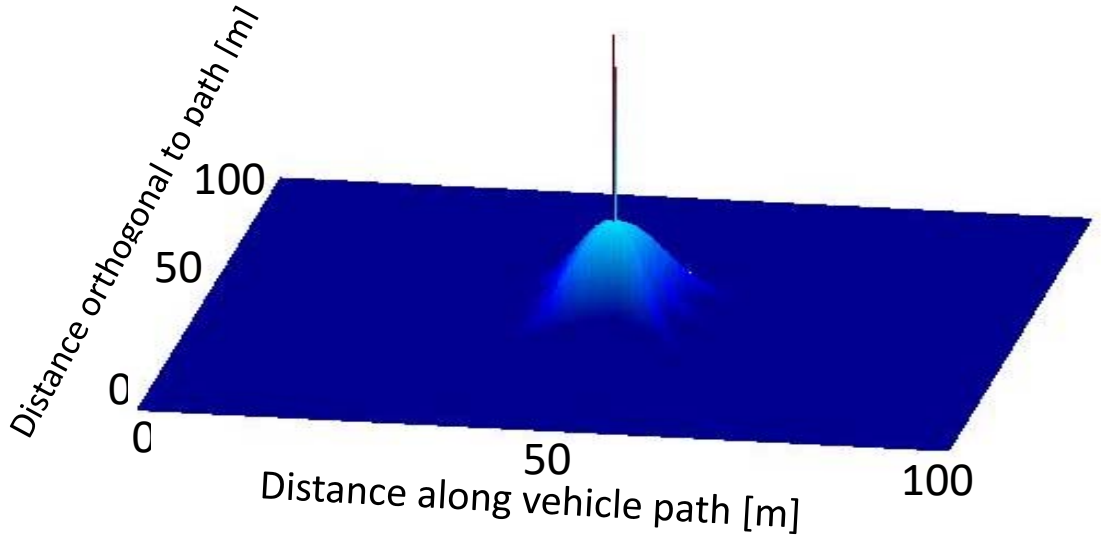


Fig. 2a. Point spread function for three observation points at 0, 50, 100 m along the vehicle path, corresponding to three adjoined basis functions from widely separated aspects.

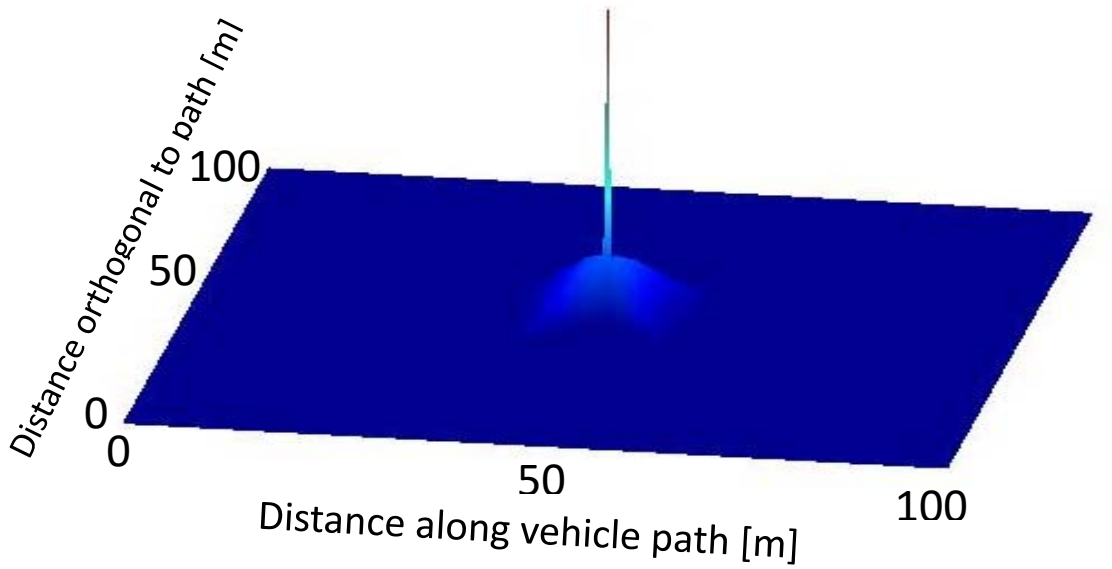


Fig. 2b. Point spread function for five observation points at 0, 25, 50, 75, 100 m along the vehicle path, corresponding to five adjoined basis functions from widely separated aspects.

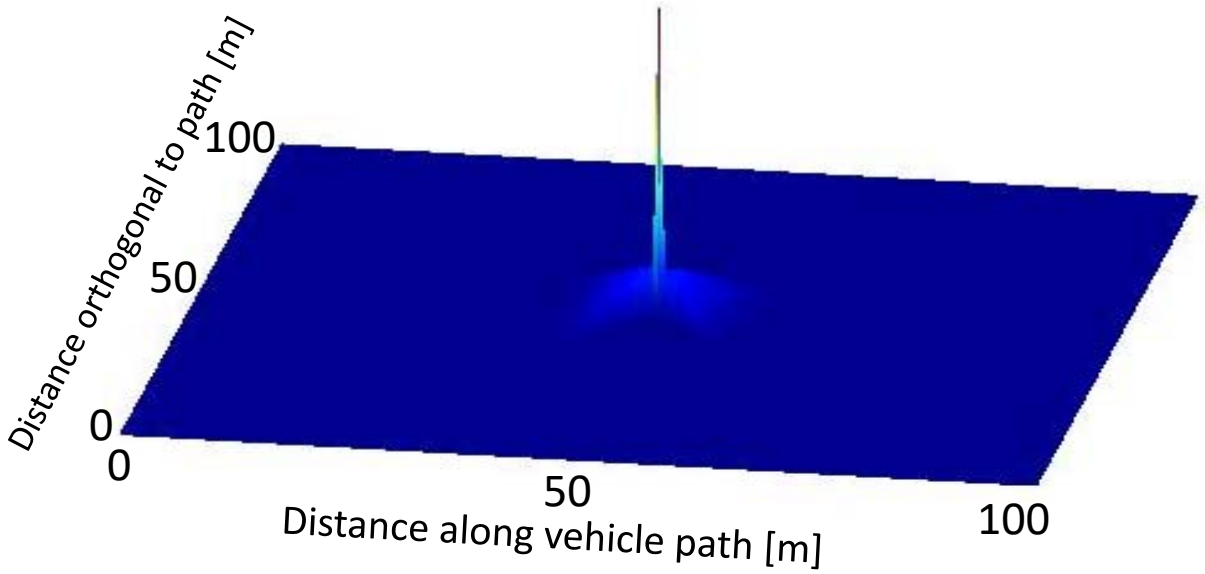


Fig. 2c. Point spread function for seven observation points at 0, 17, 33, 50, 67, 83, 100 m along the vehicle path, corresponding to seven adjoined basis functions from widely separated aspects.

The $\boldsymbol{C}^{-1}$ operation of the MMSE estimator compensates for an imperfect point spread function; the ideal version is an impulse. To implement this compensation operation, $|\rho|_{max}$ (the maximum magnitude of the basis function cross correlation coefficients in the ***C***-matrix) should be proportional to the maximum sidelobe level of the point spread function, which is 1/*N*. The largest possible value of $|\rho|_{max}$ equals one. This value occurs when *N*=1, i.e., observation from a single aspect where wide azimuth resolution causes target points at the same range to be unresolvable. The proportionality constant between 1/*N* and $|\rho|_{max}$ must therefore be unity:

$$|\rho|_{max} = 1/N \tag{17}$$

where *N* is the number of different aspects or SAS observation points used for target imaging.

Resolution

System resolution can be investigated by imaging two point-targets placed close together. The corresponding two-point target image has two basis functions, and K=2 in (1). The example in (12)-(16) is relevant to this resolution problem. Since the corresponding 2×2 ***C***-matrix has only one off-diagonal term, the MSE result in (16) can be used with

$$|\rho| = |\rho|_{max} = 1/N \tag{18}$$

where *N* is the number of target observations from different aspects. From (16) and (18), the mean-square error for estimating the reflectivities of two closely spaced target points using the complete MMSE estimator with $\boldsymbol{C}^{-1}$ sharpening and N different aspects or SAS observation points is

$$\begin{aligned} MSE_{MMSE} = \frac{N^2(N^2+1)}{(N^2-1)^2}\,\sigma_n^{\,2} &= 2.22\,\sigma_n^{\,2} \text{ for } N=2 \\ &= 1.41\,\sigma_n^{\,2} \text{ for } N=3 \\ &= 1.13\,\sigma_n^{\,2} \text{ for } N=5 \\ &= 1.06\,\sigma_n^{\,2} \text{ for } N=7 \end{aligned} \tag{19}$$

where $\sigma_n^{\,2}/2$ is the expected power of the noise samples that are added to the elements of the ***r***-vector, i.e., after space-time matched filtering. The corresponding SNR can be significantly larger than the SNR corresponding to environmental noise at the matched filter inputs. For N=1, two points at the same range cannot be separated to obtain their sample values and MSE is unbounded.

Conventional SAS (maximum likelihood) processing vs MMSE estimation

The mean-square error for maximum likelihood (beam form and matched filter) SAS processing is obtained by replacing the basis function correlation matrix $\boldsymbol{C}$ by a diagonal matrix $\boldsymbol{D}$, where $\boldsymbol{D}$ is obtained by setting the off-diagonal elements of $\boldsymbol{C}$ to zero or by using an identity matrix if the basis functions are energy normalized. $\boldsymbol{D}^{-1}$ compensates for a lack of basis function energy normalization, but does not have the sharpening capability of the $\boldsymbol{C}^{-1}$-matrix. The mean-square error for the maximum likelihood estimator is

$$MSE_{ML}=(1/K)E\{(\boldsymbol{r}-\boldsymbol{a})^{T}(\boldsymbol{r}-\boldsymbol{a})^{*}\}$$

$$= (1/K)E\{[(\boldsymbol{Ca}+\boldsymbol{n})-\boldsymbol{a})]^{T}[(\boldsymbol{Ca}+\boldsymbol{n})-\boldsymbol{a}]^{*}\}$$

$$= (1/K)E\{[(\boldsymbol{C}-\boldsymbol{I})\boldsymbol{a}+\boldsymbol{n}]^{T}[(\boldsymbol{C}-\boldsymbol{I})\boldsymbol{a}+\boldsymbol{n}]]^{*}\}. \quad (20)$$

For K=2,

$$MSE_{ML}= (1/2)E\{\left[\begin{bmatrix}0 & \rho\\ \rho & 0\end{bmatrix}\begin{bmatrix}a_1\\ a_2\end{bmatrix}+\begin{bmatrix}n_1\\ n_2\end{bmatrix}\right]^{T}\left[\begin{bmatrix}0 & \rho\\ \rho & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{a}_1\\ a_2\end{bmatrix}+\begin{bmatrix}n_1\\ n_2\end{bmatrix}\right]^{*}\}. \quad (21)$$

Using (14), (15), and (18),

$$MSE_{ML} = \sigma_n^{2}+|\rho|^{2}(|a_1|^{2}+|a_2|^{2})$$

$$= \sigma_n^{2}+ (|a_1|^{2}+|a_2|^{2})/N_{ML}^{2} \quad (22)$$

where N_{ML} is the number of observations at various aspects. Since the ML estimator is used in conventional SAS, N_{ML} is the number of transmit/receive points along the vehicle path for construction of a conventional SAS image.

In the maximum likelihood estimator with $\rho\neq0$, the value of the a_1 estimate depends on a_2 and vice-versa. This mutual interference causes an increase in MSE.

The SNR after pulse compression (correlation of data with the basis functions) is

$$SNR_{ML}= (|\boldsymbol{a}_1|^{2}+|\boldsymbol{a}_2|^{2})/ \sigma_n^{2} \quad (23)$$

and

$$MSE_{ML}= \sigma_n^{2}(1+SNR_{ML}/N_{ML}^{2}). \quad (24)$$

Equations (19) and (24) can be used to demonstrate the efficacy of the minimum mean-square error (MMSE) estimator relative to the maximum likelihood (ML) version, which is the conventional SAS receiver. For a two-point target, the MMSE estimator generally requires significantly fewer multi-aspect observations *N*, where the difference depends on SNR. For targets with many more points, the MMSE estimator is expected to be even more advantageous because of interdependence between a larger number of estimated parameters (target sample values).

Before making numerical comparisons, it is important to understand the significance of the SNR measured after computation of the ***r***-vector in (5). Each component of the ***r***-vector is computed by forming the inner product of the data vector with a conjugated basis vector. Each of the basis vectors is formed by a raster scan of delayed signal time samples from a particular target point received by vehicle array element #1, followed by delayed signal time samples from element #2, etc. The resulting number of space-time samples can be very large if the signal is modulated to increase its time-bandwidth product.

The noise-free data vector is modeled by a weighted sum of the basis vectors, where the weights are the unknown complex values of the target sample points. The correlation process in (5) forms a weighted sum of the data space-time samples, with different weights for different target points. If independent noise is added to each sample, then the noise variance after summation equals the sum of the noise variances of individual samples before summation, and the signal amplitude after summation equals the sum of the signal sample amplitudes before summation. The signal power after summation is the square of the corresponding amplitude, and the ratio of signal power to noise variance after summation (the internal SNR) is the external SNR multiplied by the number of samples in the space-time data or individual basis vectors.

For multi-pulse coherent processing, the data vector from a new observation is adjoined to the data vector from previous observations, and the basis vector corresponding to a given target point from the new observation is adjoined to the previous basis vector corresponding to the same target point. The resulting number of samples in the data vector or individual basis vectors equals the product of the number of delayed time samples corresponding to the signal, the number of physical array elements on the vehicle, and the number of independent multi-aspect observations *N*. Even for a relatively small physical array with (say) 11 elements, a relatively short signal with (say) 100 independent complex time samples, and only 3 observations from different aspects, the internal SNR associated with noise added

to the ***r***-vector is the external (environmental) SNR multiplied by 3300 or increased by 35 dB. Conversely, the externally measured SNR (average input signal power divided by average environmental noise power) is smaller than the internal SNR by a factor of 0.0003 or -35 Db. If the echo amplitude varies, the processing gain is best represented as the ratio of total multi-element, multi-aspect echo energy to noise power, which corresponds to the usual pulse compression gain of $2E/N_0$ for phase/frequency modulated signals, where $N_0/2$ is the noise power spectral density.

The ML and MMSE estimators can be compared by setting the mean-square error of the maximum likelihood (conventional SAS) processor with N_{ML} multi-aspect observations in (24) equal to the mean-square error of the MMSE estimator with *N* observations in (19):

$$\sigma_n^2(1+SNR_{ML}/N_{ML}^2) = \frac{N^2(N^2+1)}{(N^2-1)^2}\sigma_n^2 \tag{25}$$

or

$$N_{ML} = \{SNR_{ML}/[\frac{N^2(N^2+1)}{(N^2-1)^2} - 1]\}^{1/2} \tag{26}$$

where SNR_{ML} is internal to the processor, after correlation of the data vector with *K* basis functions corresponding to *K* different target points. For the current example, *K*=2.

A MMSE estimator with *N*=3 observations corresponds to a ML processor with N_{ML} observations, where (26) indicates that

$$N_{ML} = [SNR_{ML}/0.406]^{½}. \tag{27}$$

If the internal SNR of the ML processor after pulse compression and array processing is 100 (20 dB) then N_{ML} = 16 when *N*=3. If the MMSE estimator uses *N*=5 observations from different aspects, then the corresponding ML processor with internal SNR=20 dB needs N_{ML} = 20 such observations for equivalent performance.

If the off-diagonal elements of ***C*** are set to zero, the inverse of the resulting diagonal matrix ***D*** can be used to compensate for a lack of basis function energy normalization, but $\boldsymbol{D}^{-1}$ does not have the sharpening or decorrelation capability of the full $\boldsymbol{C}^{-1}$ matrix. Standard array processing methods (including SAS) generally disregard the full MMSE solution by using ***D*** instead of ***C***, where ***D***=***I*** when basis functions are pre-normalized**.** This formulation is useful, since (1) the resulting receiver corresponds to maximum likelihood detection and still benefits from space-time correlation processing (beam forming and pulse compression), (2) the two methods are nearly identical for small $|\rho|$, which corresponds to a large number of observations (transmit/receive points) at different aspects as in conventional SAS, and (3) the dimension of the ***C*** matrix can become very large, making $\boldsymbol{C}^{-1}$ difficult to calculate. This report,

however, will demonstrate the advantages of the full MMSE formulation for fast SAS imaging via relatively few observations with spotlights and for other applications that utilize coherent multi-aspect processing. In these applications, $|\rho|$ can be reduced but often is not negligible.

Computer simulation with perfectly known vehicle locations and rotations

The simulation starts with a specification of the transmitted signal. The chosen signal is currently a linear chirp with center frequency f_0=37.5 KHz and bandwidth B=15 KHz, between a low frequency of 30 KHz and a high frequency of 45 KHz. The wavelength at the center frequency is λ=0.04 m. The nominal two-way range resolution cell (with SNR=1) is $\Delta R = c/(2B) = 0.05$m. Signal duration is T=4 ms, corresponding to a minimum range of cT/2=3m to avoid simultaneous transmission and reception. The complex-valued signal is oversampled by a factor of 4, resulting in IT_{MAX}=720 complex time samples or 180 statistically independent samples for SNR calculations.

A wide-beam transmitter is at the center of the receiving array, which consists of 11 elements spaced by the average wavelength (d=0.04m). The total receiving array length is 0.4 m. Echoes from the wide-beam transmitter are assumed to be measured at each array element and stored, so that receiving beams (spotlights) can be formed from stored data. The spotlights are steered via relative delays of the element outputs in order to focus on the same target point from different locations along the vehicle path.

The capability of the MMSE estimator is demonstrated by transmitting and receiving from only three points along the vehicle path: x=0 m, 50m, and 100m. Fig. 2 (left side) shows the point spread function corresponding to the chosen signal/array parameters and the three transmit/receive locations.

A target consisting of at least K=25 sample points in a $K^{1/2} \times K^{1/2}$ (at least 5×5) array is centered at x=50m, y=50m. The distance between target sample points is 2.5cm in both x and y directions. This distance is chosen to be half the signal range resolution cell (at SNR=1) in order to demonstrate MMSE sharpening capability in multiple directions.

The distance from the transmitter to each target point and from each receiver array element to each target point is computed for each of the three transmit/receive locations. These distances are used to compute the corresponding delays along the ray paths shown in Fig. 1.

At each transmit/receive location, the signal, delays and specified target samples are used to obtain the expected echo time samples at each of the receiver array elements from a target point with unit reflectivity at each target sample location. The collection of these delayed signal samples constitutes a partial space-time basis function corresponding to the chosen target point. This partial basis function corresponds to one of (say) three transmit/receive locations along the vehicle path.

A different partial basis function is obtained for each chosen transmit/receive location. Each partial basis function is converted into a vector of delayed signal samples by adjoining all the time samples from the second array element to all the time samples from the first array element, and continuing until time samples from all the array elements at a given transmit/receive location have been adjoined.

A multi-aspect basis vector $\boldsymbol{b}_K$ for target point k is obtained by adjoining the partial basis vectors from all relevant (e.g., three) transmit/receive locations. There are K such vectors altogether, one for each target point. The $\boldsymbol{b}_K$ vectors are used to construct the $\boldsymbol{C}$-matrix for MMSE estimation of the $\boldsymbol{a}$-vector containing the K target samples as in (6). The dimension of the $\boldsymbol{C}$-matrix is $K \times K$.

A data vector corresponding to target point k with multi-aspect basis vector $\boldsymbol{b}_K$ is constructed by multiplying all the elements of $\boldsymbol{b}_K$ by a simulated complex target sample a_k and adding an independent noise sample to each element.

The complete data vector corresponding to (1) is the sum of the data vectors corresponding to each of the target points, and represents a space-time echo from the whole target, as observed at all specified locations along the vehicle path.

The inverse of the $\boldsymbol{C}$-matrix is computed with a customized complex version of subroutine MATINV from Numerical Recipes [6]. For a 5×5 target image, $\boldsymbol{C}$ and $\boldsymbol{C}^{-1}$ have dimension 25×25.

ML and MMSE target sample estimates of a 5×5 one-point image for various input (environment) SNR values are shown in Fig. 3. Larger, multi-point images are in Fig. 4.

The imaging capability illustrated by the examples is impressive, but it will be shown that this performance depends on very accurate measurement of transmitter/receiver location and orientation. If the required measurement accuracy is unavailable, high-performance MMSE imaging depends on an image-based correction method that can compensate for small location/rotation measurement errors.

Computer simulation with imperfectly known vehicle locations and rotations

The multi-aspect space-time basis functions depend on delays between the transmitter, various target points, and receiver array elements at transmit/receive points along the vehicle path. These delays are sensitive to differences between measured vehicle locations/rotations and their true values. Degraded ML and MMSE images may contain information about the measurement errors that caused the degradation. This information ideally could be used to mitigate the effects of the errors. A sharper image is expected when the measured location/ rotation parameters are closer to their true values, but it is not certain that an image sharpness measure can provide reliable corrections in a gradient search optimization process. Since the $\boldsymbol{C}$-matrix depends strongly on measured location/rotation, the MMSE image is more strongly affected than the ML image by location/rotation errors. This sensitivity difference provides additional information for error correction.

To appreciate the effects of location/position error, a one-point target sample distribution is used to investigate a simple error in the computer simulation: The actual transmitter/receiver is 0.5 cm further along the vehicle path at the first observation point than measurements indicate. Other errors at all observation points are set to zero. The resulting images at SNR=1 (0 dB) are shown in Fig.5. The error-free versions are in Fig. 3c.

The ML image in Fig. 5 is much less sensitive to vehicle location error than the MMSE image. Without Fig. 3, an observer would not know the shape of the error-free MMSE image from inspection of Fig. 5.

Fig. 3a. True sample values of a one-point target located 50m from the vehicle path as in Fig.2. The samples are 2.5 cm apart in both directions (half of a range resolution cell at SNR=1).

.000	.000	.000	.000	.000
.000	.000	.000	.000	.000
.000	.000	1.000	.000	.000
.000	.000	.000	.000	.000
.000	.000	.000	.000	.000

Fig. 3b. Amplitude normalized target sample estimates without noise or unknown transmitter/receiver displacements, using echoes from three transmit/receive locations (Fig. 2a). Left: Maximum likelihood (ML) estimate of the target samples in Fig. 3a. The ML estimate uses only the diagonal elements of the ***C***-matrix (the multi-aspect basis function correlation matrix), which implements basis function normalization without sharpening. Right: MMSE estimate using the multi-aspect basis normalization and sharpening capability of the full ***C***-matrix. Samples are 2.5 cm apart. Range varies between 50m and 71 m.

.354	.231	.225	.231	.354
.453	.306	.360	.306	.453
.245	.635	1.000	.635	.245
.487	.310	.347	.310	.487
.331	.213	.225	.213	.331

.000	.000	.000	.000	.000
.000	.000	.000	.000	.000
.000	.000	1.000	.000	.000
.000	.000	.000	.000	.000
.000	.000	.000	.000	.000

Fig. 3c. Amplitude normalized target sample estimates with external SNR (ratio of average signal power to average noise power at the receiver input) equal to 1 (zero dB). Transmitter/receiver locations/rotations are known exactly. Both estimators use echoes from three transmit/receive locations, as shown in Fig. 2a. Left: Maximum likelihood (ML) estimate of the target samples in Fig. 3a. Right: MMSE estimate using the full ***C***-matrix. Samples are 2.5 cm apart. Range varies between 50m and 71 m.

.356	.227	.227	.231	.359
.457	.303	.357	.307	.459
.244	.627	1.000	.638	.242
.485	.313	.340	.299	.482
.333	.225	.226	.218	.330

.033	.047	.036	.028	.027
.038	.007	.019	.041	.036
.010	.061	1.000	.064	.045
.028	.064	.075	.043	.031
.023	.031	.028	.039	.013

Fig. 3d. Amplitude normalized target sample estimates with external SNR (ratio of average signal power to average noise power at the receiver input) equal to 0.1 (-10 dB). Transmitter/receiver displacements are known exactly. Both estimators use echoes from three transmit/receive locations, as shown in Fig. 2a. Left: Maximum likelihood (ML) estimate of the target samples in Fig. 3a. Right: MMSE estimate using the full C-matrix. Samples are 2.5 cm apart. Range varies between 50m and 71 m.

.362	.218	.230	.232	.369
.467	.295	.353	.309	.470
.242	.608	1.000	.644	.236
.482	.318	.327	.274	.473
.339	.251	.228	.228	.331

.108	.155	.119	.091	.088
.126	.023	.062	.133	.117
.034	.199	1.000	.209	.147
.093	.210	.245	.141	.101
.077	.100	.090	.128	.042

Fig. 4. True target images (left), ML image estimates (center), and MMSE estimates (right) using three aspects as shown in Fig. 2a. Input SNR=10 (10dB). Image samples are 2.5 cm apart. Range is 50m to 71m. Vehicle location/rotation parameters are known exactly.

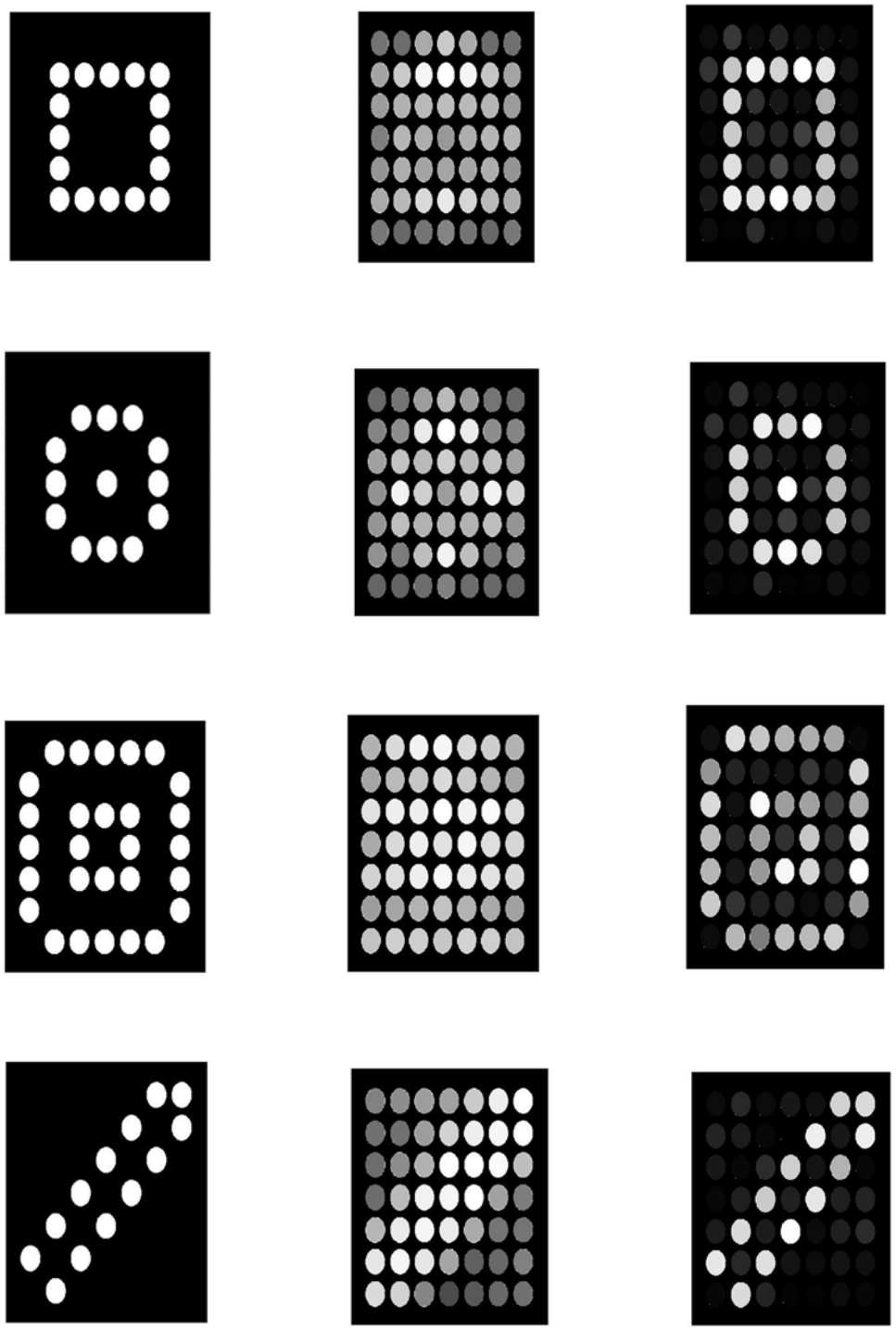


Fig. 5. ML (left) and MMSE (right) images of a one-point target at SNR=1 (0dB) as in Fig. 3c, but with an uncompensated 0.5 cm translation of the vehicle along its path at the first of three observation points.

.376	.250	.266	.259	.363		.023	.459	.317	.716	.156
.490	.296	.322	.218	.406		.395	.832	.769	.469	.358
.219	.571	1.000	.750	.321		.245	.468	1.000	.773	.447
.501	.344	.372	.385	.522		.181	.207	.449	.212	.094
.341	.211	.219	.228	.342		.128	.254	.145	.117	.073

Since the ML image in Fig. 5 is not very different from the ML image in Fig. 3c, what error magnitude causes the ML image to degrade significantly? To answer this question, the position error was increased from 0.5 cm to 1 cm. Fig. 6 shows that this increase is enough to significantly degrade an ML image constructed from three observations. More observation points increase the likelihood of location/rotation errors, but averaging may result in better images.

Fig.6. ML (left) and MMSE (right) images of a one-point target at SNR=1 (0dB) as in Fig. 3c, but with an uncompensated 1 cm translation of the vehicle along its path at the first of three observation points.

.422	.350	.187	.411	.411		.136	.282	.028	.389	.070
.764	.710	.203	.419	.050		.118	.706	.694	.590	.248
.538	.505	.770	1.000	.583		.239	.422	1.000	.842	.651
.485	.610	.436	.691	.652		.271	.260	.961	.326	.435
.494	.353	.125	.415	.371		.515	.771	.162	.608	.478

The above experiments imply that very small perturbations of location/rotation parameters can have large consequences. A gradient optimization process for location/rotation error correction should estimate the gradients of an image quality measure using very small location/rotation differences, and the constant that multiplies the gradients should be small to limit step size.

Since the MMSE image is very sensitive to location/rotation error, perhaps it can be used to correct such errors if appropriate MMSE image features (e.g., sharpness measures) are used. The different error sensitivity of ML and MMSE images provides another feature design opportunity; measures of the disparity or similarity of the ML and MMSE images may be useful for location/rotation error correction.

An important criterion for choosing image features for error correction is that the change of the image feature values for a small parameter change in a location/rotation model should reliably indicate the direction of change of the investigated parameter in order to better model or correct the parameter. This criterion should apply even when current estimates of the other parameters are erroneous. If an image feature that satisfies this criterion exists, then it can be used in a gradient optimization procedure for error correction.

Many image features have been tested, but none has consistently satisfied the above reliable gradient criterion for MMSE location/rotation error correction. Some of the tested image features are described in the following paragraphs.

Minimum image entropy was considered as a focus criterion but was rejected because it tends to make location/rotation corrections such that the estimated image is concentrated at a single point, even if the true target image is different.

To understand this result, assume that the image consists of K unit-amplitude pulses, with all other image samples set to zero. This image is very sharp and well-focused, since the pulses are not smeared. The entropy of the pattern is obtained by first normalizing the sum of the pulses to one, resulting in K pulses with amplitude $(1/K)$ and all other samples equal to zero. The entropy of the normalized pattern is

$$H= -\sum_{k=1}^{K} p_k log(p_k) \quad (28)$$

where $p_k = 1/K$. The zero-valued points are neglected because as ϵ approaches zero, $\epsilon \log(\epsilon)$ also approaches zero. Substituting $p_k = 1/K$ into (28),

$$H=-(1/K)[K \log(1/K)]=\log(K). \quad (29)$$

Since $\log(K)$ is a monotone increasing function of K, the entropy gets larger as the number of pulses increases. H is minimized when K=1, which corresponds to a one-pulse image.

Various sharpness measures were considered, with best results for the following version:

(1) The $\hat{\boldsymbol{\alpha}}$-vector of estimated complex target samples is converted into a 2-D image.
(2) For each image sample, the magnitude of the difference between the sample and each of its nearest neighbors is computed.
(3) The maximum nearest neighbor difference magnitude for each sample is obtained.
(4) The M largest maximum nearest neighbor differences for all image samples are found and their average is calculated. Good performance was obtained for M=7, but this parameter can be adjusted as required.
(5) The resulting sharpness measure is divided by the average magnitude of the $\hat{\boldsymbol{\alpha}}$-vector components for normalization.

An alignment measure was used to quantify translation and/or non-uniform sample weighting of a MMSE image relative to the corresponding ML image. These effects are caused by the extra sensitivity of MMSE estimates to vehicle location/rotation errors, relative to ML estimates. The alignment measure is computed as follows:

(1) The ML and MMSE vectors containing estimated complex target samples are converted into 2-D real-valued images of sample magnitudes, f(x,y|ML) and f(x,y|MMSE).
(2) Each of the 2-D images is normalized by dividing each sample magnitude by the average of all sample magnitudes in the image.
(3) For each normalized image, the first moments in the x and y directions are computed, e.g., E(x|ML)=sum over x and y of x[f(x,y|ML)] and E(y|ML)=sum over x and y of y[f(x,y|ML)].
(4) The differences between first moments are squared and summed:

$$d^2=[E(x|MMSE)-E(x|ML)]^2 + [E(y|MMSE)-E(y|ML)]^2 . \quad (30)$$

(5) The alignment measure is exp(-αd), where the constant α is used to control sensitivity. Perfect alignment (as measured by first moment differences) corresponds to an alignment measure of one; the measure decreases as first moment mismatch increases.

A composite measure consisting of a weighted sum of sharpness and alignment measures was used to estimate location/rotation errors. The composite measure was incrementally maximized by choosing a location/rotation adjustment proportional to the gradient of the measure. This gradient is the change in the measure value divided by a very small increase in an estimated location/rotation parameter that causes the change. A small adjustment of each parameter was made in turn, followed by another sequence of small adjustments, until no further adjustments were required (convergence of the gradient ascent procedure).

This process did not adequately compensate simulated location/rotation errors because the gradients did not consistently indicate the correct sign (positive or negative) for required parameter changes. Additional and/or better image features are needed for reliable image-based compensation of vehicle location/rotation errors. As in the alignment measure, the utilization of two images (ML and MMSE) with different error sensitivities should be helpful.

Three-dimensional imaging

Generalization to 3-D imaging is straightforward:

(1) The sparse sampling permitted by MMSE imaging allows time for the vehicle to move out of the range-azimuth plane and back into the plane (if necessary) for the next in-plane observation. The vehicle can weave up and down through the range-azimuth plane.
(2) Extra transmit/receive operations are performed while the vehicle is outside the range-azimuth plane (at nonzero elevations).
(3) The k^{th} basis vector is extended to include an extra out-of-plane observation by adjoining delayed signal time samples from a raster scan of the receiver array elements to the previous version of the k^{th} basis vector. The new signal delays correspond to a new, out-of-plane echo from the k^{th} target sample point, which is now part of a cube rather than a square.

A zig-zag target approach trajectory may look more like a corkscrew pattern if 3-D imaging is used.

Consequences of MMSE estimation with non-orthogonal basis functions

MMSE estimators generally avoid non-orthogonal basis functions because there is no well-defined limit to the number of such basis functions that should be used (lack of completeness), resulting in extra design constraints or specifications [7], e.g., the order of an autoregressive predictor [8]. In the present case, the number of basis functions equals the number of estimated target samples, which depends on the sample density (resolution) and size of the target image. In a cluttered environment, the minimum image size is determined by the spotlight width, since the image can be adjusted to include fewer samples in the range direction and more samples in azimuth. Assuming uniform sample spacing in range and azimuth, the sample density is related to the range resolution cell ΔR, which varies inversely with signal bandwidth and the square root of SNR.

Even if the basis functions were orthogonal rather than non-orthogonal, the same considerations as above would apply if the number of basis functions equals the number of estimated target samples.

The consequences of specifying the density of estimated target samples {and thus the number of basis functions for a given target size) are illustrated in Fig. 3, which illustrates SAS simulation results for a sample spacing of $\Delta R/2$, where ΔR corresponds to SNR=1. Accurate MMSE estimation with three transmit/receive locations is obtained for SNR $\geq$ 1, but accuracy is lost for a smaller SNR. Denser sampling could have been used to take advantage of better resolution for SNR>1, and sparser sampling could have been used for better efficiency when SNR<1.

Summary and conclusion

Synthetic aperture sonar (SAS) target imaging can be formulated as a minimum mean-square error (MMSE) estimation problem instead of the usual maximum likelihood (ML) approach. Each basis function for MMSE target sample estimation is associated with a different target point and is described by predicted echoes received from that point. Multi-aspect SAS data are modeled as a weighted sum of these basis functions, with weights corresponding to complex samples of the target image.

The first part of the MMSE estimator involves correlation of data with conjugated basis functions, and corresponds to a conventional ML receiver or space-time SAS processor. The second part is a sharpening operation performed with the inverse of the basis function correlation matrix. Implementation of both parts of the MMSE receiver allows efficient imaging with only a few (e.g., three) transmit/receive (T/R) operations at different aspects, and with a relatively small physical array (0.4 m at 37.5 kHz center frequency). If the vehicle position is known at each T/R location, the T/R points are not constrained to lie on a straight-line vehicle path. The path should allow for observations at different aspects, e.g., weaving above and below the range-azimuth plane for 3-D SAS imaging. Excessive input signal-to-noise ratio (SNR) is not required, and low input SNR can be compensated by using a signal with more independent time samples, a larger receiving array, or more T/R observations.

Impressive multi-aspect MMSE target images from only three T/R locations requires (1) highly accurate vehicle location/rotation measurements with errors less than 5mm or (2) error correction via image quality assessment. Extra image quality features are obtained by comparing ML and MMSE images,

since MMSE is more sensitive than ML to location/rotation errors. Some promising features have been found, but currently they are insufficient to solve the image-based error compensation problem. Some motion-based errors can be minimized by using a Doppler tolerant signal with no range-Doppler coupling [9].

MMSE multi-aspect imaging has some important applications in addition to SAS. A low-risk application is to radar/sonar imaging from several fixed T/R platforms or a single fixed transmitter and several fixed receivers (multi-static imaging). The risk is low because fixed transmitter and array element locations can be accurately measured. Surface ships with locations/rotations that can be measured with lasers also provide good potential platforms for multi-aspect MMSE radar/sonar imaging.

If receivers are already in place at the three locations in Fig. 2a, then one transmission from (say) the center location is sufficient to generate a MMSE target image estimate. Subsequent transmissions can be used to scan the environment, for tracking, or to monitor changes in a target image. This one-pulse imaging capability is important for many applications, including underwater intruder identification with fixed array positions.

References

[1] W.G. Carrara, R.S. Goodman, and R.M. Majewski, *Spotlight Synthetic Aperture Radar* (Artech House, 1995).

[2] M.I. Skolnik, *Radar Handbook* (McGraw-Hill, 1970) p. 4-8.

[3] A.K. Jain, *Fundamentals of Digital Image Processing* (Prentice Hall, 1989) pp. 431-470.

[4] L.L. Scharf, *Statistical Signal Processing* (Addison-Wesley, 1991) p. 239.

[5] C.L. Lawson and R.J. Hanson, *Solving Least Squares Problems* (Prentice-Hall, 1974) p. 29

[6] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vetterling, *Numerical Recipes* (Cambridge Univ. Press, 1986).

[7] H.L. Van Trees, *Detection, Estimation, and Modulation Theory* (Wiley, 1968).

[8] S.L. Marple, Jr., *Digital Spectral Analysis* (Prentice-Hall, 1987)

[9] R.A. Altes and E.L. Titlebaum, *Bat Signals as Optimally Doppler Tolerant Waveforms*, J. Acous. Soc. Am. 48, 1014 (1970).