Target separation in SAR image with the MUSIC algorithm

Philip Thompson
Technological Institute of Aeronautics ITA
Department of Electronics
São José dos Campos, Brasil
e-mail: philipthomp@gmail.com

Matteo Nannini and Rolf Scheiber
German Aerospace Centre - DLR
Microwaves and Radar Institute
Oberpfaffenhofen, Germany
e-mail: matteo.nannini@dlr.de

***Abstract*—The aim of this work is to exploit the MUSIC algorithm performance in order to enhance target separability in range and azimuth, i.e. achieve point targets separation inside a resolution cell. Simulations have been done in order to plan and check the feasibility of a super-resolution experiment that took place in September 2006 on the test site of Oberpfaffenhofen (Germany). The data set has been acquired with the E-SAR system of the DLR in X-Band. The targets to be separated were seven small corner reflectors that have been placed in a way that their response falls in one or, at maximum, two resolution cells of the standard Fourier SAR image. A post-processing implementation of the MUSIC algorithm has been proposed allowing, in the already focused SAR image, to retrieve the targets geometry. Conditions and analysis of the results have been carried out.**

I. Introduction

The most important advantage of SAR systems is the high azimuth resolution achieved due to the presence of the synthetic aperture. The range resolution is defined by the hardware characteristics of the SAR sensor. A standard approach for obtaining SAR images is the Fourier based Matched Filtering. The most intuitive way in order to enhance the range resolution, is to operate with a larger bandwidth either within the same pulse, alternating the carrier chirp frequency as it is done in the stepped frequency mode or exploiting the interferometric wavenumber shift principle using two different acquisitions with slightly different viewing angles. For the azimuth resolution, a Fourier based enhancement could be obtained by steering the antenna (Spotlight) or by repeated acquisitions with different squint angles. In the recent years an alternative way that makes use of modern spectral estimation techniques like MUSIC has been approached with raw data in spotlight acquisitions [1], [4].

The MUSIC algorithm was initially proposed to estimate the Direction Of Arrival (DOA) of signals [5]. This can be done with time measurements of the e.m. field emitted by several sources with the help of an array of sensors. The application of this method in SAR imaging processing requires some adaptation and, for this reason, two main drawbacks appear. The first is intrinsic in the SAR nature: the MUSIC algorithm exploits the signal statistics (i.e. the data covariance matrix) in order to build up the noise subspace and to exploit its characteristics to retrieve the target location. To generate the statistics of interest, multiple acquisitions are required, therefore a single SAR survey is not an ideal working condition for it. Anyhow, there are methods in the literature which transform the single acquisiction into a multiple one. The second reason is that, the complex reflectivity information can not be easily recovered. The MUSIC response is not proportionally related to the target backscattering power. If a target is detected, the algorithm response consists in a high isolated power peak corresponding to its position, but this power is a measure of confidence rather than of target backscattering. For these reasons a 2D SAR image with higher quality can not be achieved by the help of MUSIC alone.

Although it has been shown that this method is not well suited for a complete 2D super-resolution SAR image, it is still worth to make an investigation in a more particular application field related to target separation. A post-processing implementation of the MUSIC algorithm based on [2] is proposed allowing, in the already focused SAR image, to retrieve the targets geometry and/or the separation of different scattering centers.

II. The MUSIC algorithm for radar imaging

A. MUSIC data model

In a general context of application, the MUSIC model is seen as an estimation problem where D vector parameters $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_D$ (range, azimuth position coordinates) of D objects of interest (backscatterers) have to be estimated. The data acquisition model is:

$$\mathbf{X}(t) = \mathbf{A}\mathbf{F}(t) + \mathbf{W}(t) \qquad t = 1, \ldots, K \tag{1}$$

where the data-vector $\mathbf{X}(t) = [X_1(t) \cdots X_L(t)]^T$ is acquired K times and L is the number of spatial samples. The noise-vector is represented by $\mathbf{W}(t) = [W_1(t) \cdots W_L(t)]^T$ and $\mathbf{F}(t) = [F_1(t) \cdots F_D(t)]^T$ is the vector signal associated to the D objects. The $L \times D$ matrix $\mathbf{A}$ is defined as $\mathbf{A} = [\mathbf{a}(\boldsymbol{\theta}_1) \cdots \mathbf{a}(\boldsymbol{\theta}_D)]$, where $\mathbf{a}(\boldsymbol{\theta})$ is the so-called *steering vector function*.

The noise is assumed to have covariance matrix $E[\mathbf{W}\mathbf{W}^*] = \sigma^2\mathbf{I}$ and the vectors $\mathbf{W}$ and $\mathbf{F}$ to be decorrelated and with zero-mean. Three main assumptions are done:

- The vectors $\mathbf{a}(\boldsymbol{\theta}_1)\ldots\mathbf{a}(\boldsymbol{\theta}_D)$ (the *mode vectors*) are linearly independent.
- the covariance matrix $\mathbf{P} = E[\mathbf{F}\mathbf{F}^*]$ of $\mathbf{F}$ is non-singular.
- $L > D$.

Now, defining the data covariance matrix as $\mathbf{S} = E[\mathbf{X}\mathbf{X}^*]$, two complementary and orthogonal subspaces of $\mathbb{C}^L$ can be constructed, one of dimension D (spanned by the mode vectors) called Signal Subspace and the other of dimension $N = L - D$ called Noise Subspace. Organising the base vectors of the noise subspace in the columns of a matrix $\mathbf{E}_N$, it can be shown that the function of $\boldsymbol{\theta}$:

$$P_{MU}(\boldsymbol{\theta}) = \frac{1}{\mathbf{a}^*(\boldsymbol{\theta})\mathbf{E}_N\mathbf{E}_N^*\mathbf{a}(\boldsymbol{\theta})} \tag{2}$$

achieves peaks only when the argument $\boldsymbol{\theta} = \boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_D$, detects the source position. In a realistic case, the acquisition number is finite, therefore the covariance matrix has to be estimated by:

$$\widehat{\mathbf{S}} = \frac{1}{K}\sum_{t=1}^{K}\mathbf{X}(t)\mathbf{X}^*(t). \tag{3}$$

In order to apply the MUSIC algorithm with the estimated covariance matrix (3), an priori estimation of the $\widehat{D}$ *number of targets* D is necessary [7]. In the specific 2D radar or SAR imaging context, it is possible in some cases to achieve the image data as in Eq.(1) in such a way that the estimation parameter is $\boldsymbol{\theta} = (\omega_r, \omega_a)$ where ω_r and ω_a are respectively linearly related to the range and azimuth image coordinates. The desired parameters are $\boldsymbol{\theta}_1 = (\omega_{r1}, \omega_{a1}), \ldots, \boldsymbol{\theta}_D = (\omega_{rD}, \omega_{aD})$. To accomplish this, the data image model is assumed to be in the form of a two-dimensional complex sinusoid signal, i.e.:

$$X_{mn} = \sum_{p=1}^{D}\gamma_p e^{j\omega_{rp}m}e^{j\omega_{ap}n} + W_{mn} \qquad \begin{array}{l} m = 0, \ldots, M-1 \\ n = 0, \ldots, N-1 \end{array} \tag{4}$$

where, for the scatterer p the range frequency is ω_{rp} and azimuth frequency is ω_{ap}.

In order to use the model Eq.(1) it has to be noticed that in the SAR the data are two-dimensional and, more important, only one acquisition is done (one pass), therefore $k = 1$. These drawbacks can be overcame using the *spatial smoothing* [1], [3], [4] method (depicted in Fig.1). It consists in defining a window of dimensions $m_1 < M$ and $m_2 < N$ smaller than the image and scan it in all possible positions. For each scanning position, the submatrix defined is treated as an acquisition and its elements are raster scanned in a defined order (in our case each line is allocated after the other) into the one-dimensional vector $\mathbf{X}(t)$. Due to the linearity of the exponential, the image model in Eq.(4), can be transformed as in Eq.(1) with $K = (M - m_2 + 1)(N - m_1 + 1)$ acquisitions and with a one dimensional steering vector:

$$\mathbf{a}(\omega_r, \omega_a) = \\ [1 \quad e^{j\omega_a} \quad \cdots \quad e^{j(m_2-1)\omega_a} \\ e^{j\omega_r} \quad e^{j\omega_r}e^{j\omega_a} \quad \cdots \quad e^{j\omega_r}e^{j(m_2-1)\omega_a} \\ e^{j(m_1-1)\omega_r} \quad e^{j(m_1-1)\omega_r}e^{j\omega_a} \quad \cdots \quad e^{j(m_1-1)\omega_r}e^{j(m_2-1)\omega_a}]^T$$

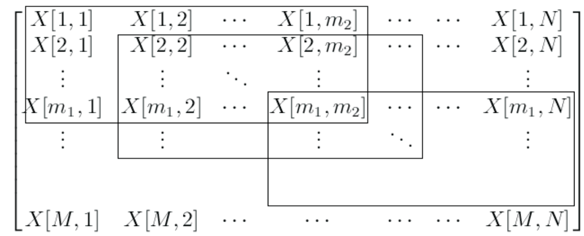


Fig. 1. Spatial smoothing over a two dimensional data array.

B. MUSIC Pre-processing

In [1], [4] the application of MUSIC for SAR image processing purpose has been carried out for the retangular formatted spotlight raw data acquisition where, after interpolation the expression of Eq.(4) can be assumed. In this work a processing scheme in order to achieve the model of Eq.(4) in an already focused 2D SAR image is described. The purpose is to avoid raw data complication features (range migration and motion compensation) which are already build into a standard SAR processor. The application of the MUSIC algorithm have been limited to a small region of the image, where a limited number of strong targets are located.

For the sake of simplicity the time continuous case will now be considered. Each point target p located at the image data coordinates (τ_p, t_p), related to range and azimuth respectively, is focused in the SAR image as a two-dimensional sinc-function:

$$y(\tau, t) = \sum_{p=1}^{D}\gamma_p \mathrm{sinc}\left[B_{sr}(\tau - \tau_p)\right]\mathrm{sinc}\left[B_{sa}(t - t_p)\right], \tag{5}$$

where the amplitude γ_p is the target reflectivity, B_{sr} is the bandwidth of the range-sinc and B_{sa} is the bandwidth of the azimuth-sinc. After selectioning the part of the image where the targets are located, the target area is windowed. To achieve the exponential form of Eq.(4) a convolution of the focused image with a predefined two-dimensional chirp such as

$$c(\tau, t) = e^{j\alpha_r\tau^2}\mathrm{rect}[\Delta_{cr}\tau]e^{j\alpha_a t^2}\mathrm{rect}[\Delta_{ca}t] \tag{6}$$

is carried out. In (6) α_r and Δ_{cr} are respectively the range-chirp rate and length and α_a and Δ_{ca} are respectively the azimuth-chirp rate and length. The result of the convolution is the sum of D chirps with length, rate and bandwidth of (6) but centered at each target image coordinates:

$$s(\tau, t) = \sum_{p=1}^{D}\gamma_p e^{j\alpha_r(\tau-\tau_p)^2}e^{j\alpha_a(t-t_p)^2}. \tag{7}$$

In the case of the chirp, this also is achieved if its bandwidth is higher than the target one due to its frequency vs. time proportionality (at the cost of a lower pulse duration). This procedure is known as data *expansion*. It has to be observed that due to the limited pulse duration of the chirp, after the convolution, the chirps centered at the position of the targets have a region of intersection which it is not reported in (7). Anyhow, these inaccuracy can be neglected because, for the application of interest, the targets are supposed close enough.

Through (7), the model (4) can be achieved if the quadratic exponential phase of the chirp is transformed into a linear exponential phase (sinosoid). This can be done trought the well known *deramping* process which simply multiplies the signal by its delayed and conjugated version (deramping signal), in our case:

$$\mathrm{der}(\tau, t) = e^{j\alpha_r(\tau-\tau_{ref})^2} e^{j\alpha_a(t-t_{ref})^2} \tag{8}$$

where τ_{ref} and t_{ref} are the defined delays. The deramping signal in (8) must cover all the time duration of the signal (7). The deramped signal can be written as:

$$X(\tau, t) = \mathrm{der}(\tau, t)s^*(\tau, t) = \sum_{p=1}^{D} \varsigma_p e^{j\nu_{rp}\tau} e^{j\nu_{ap}t} \tag{9}$$

with $\nu_{rp} = 2\alpha_r(\tau_p - \tau_{ref})$ and $\nu_{ap} = 2\alpha_a(t_p - t_{ref})$.

Defining now F_s the sampling rate in range and PRF the one in azimuth, we have that $t_{sr} = 1/F_s$ and $t_{sa} = 1/\mathrm{PRF}$ are the correspondent sampling periods. The discrete version of the model (see ((4))) will be computed as $\omega_{rp} = 2\alpha_r(m_p - m_{ref})t_{sr}^2$ and $\omega_{ap} = 2\alpha_a(n_p - n_{ref})t_{sa}^2$ where, $m_p = \tau_p/t_{sr}$ and $n_p = t_p/t_{sa}$ are respectively the sampled range and azimuth data coordinates of the target p and $m_{ref} = \tau_{ref}/t_{sr}$ and $n_{ref} = t_{ref}/t_{sa}$. The domain of search is $\{2\alpha_r(0-m_{ref})t_{sr}^2, \ldots, 2\alpha_r((M-1)-m_{ref})t_{sr}^2\}$ for range frequency and $\{2\alpha_a(0-n_{ref})t_{sa}^2, \ldots, 2\alpha_a((N-1)-n_{ref})t_{sa}^2\}$ for the azimuth one, correspoding to the selected sub-image.

III. Two simulation results

Two simulation results, with additive noise, separated by sub-resolution distances are presented. In Fig.2 the azimuth separation is in the order of 0.35 of the Fourier azimuth resolution. In Fig.3 a second result is presented for a range separation of 0.35 of the Fourier range resolution. In each case, it is shown also the percentual relation between the image size at the processing stage of Eq.(4) and the size of the submatrix chosen in the spatial smoothing method. The notation 'range×azimuth' will always be used. The profiles in the left correspond to range and in the right to azimuth. From up to bottom, the profiles correspond respectively to the first target, the second target, the sum of the two and, finally, the profile given at the output of the proposed processing. In both cases the method is able to separate with resonable accuracy the two targets.

IV. Real data

Seven small corner reflectors were placed in a test field in Oberpfaffenhofen (Germany) in the geometry shown in Fig.4(b). The radar data was acquired on September 2006 by the E-SAR system of DLR at X-band in the HH polarization. The nominal resolutions are 0.6m in azimuth and 2.12m in range. A selection of an area of 21×21 pixels in

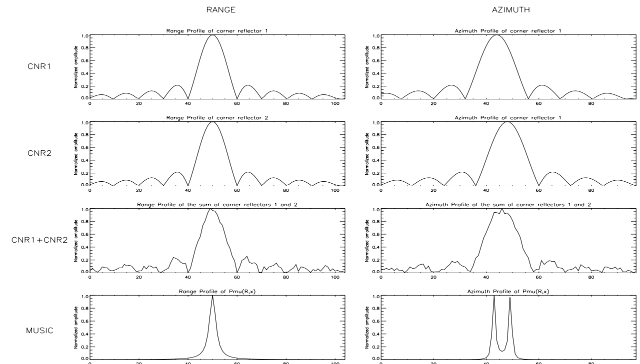


Fig. 2. Profile of simulated targets in azimuth separation. Subarray size: $10\% \times 30\%$

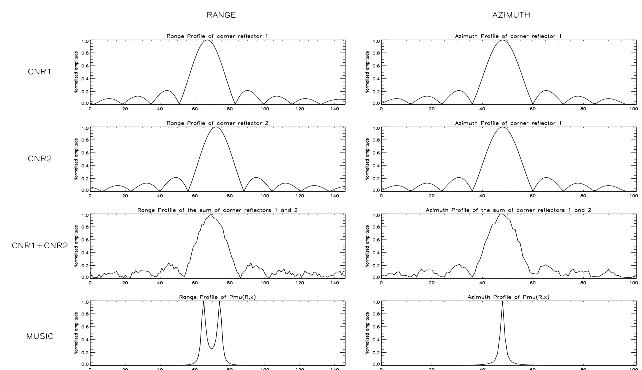
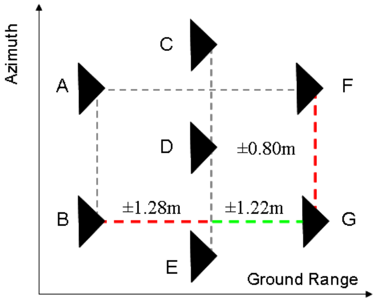


Fig. 3. Profile of simulated targets in range separation. Subarray size: $30\% \times 10\%$.

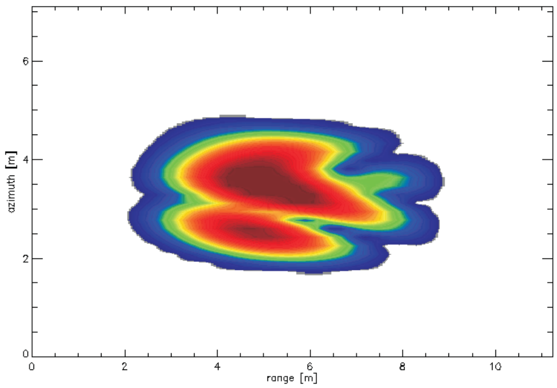
the surrounding region of the corners image was done. The subarray size was around $25\% \times 20\%$. Fig.4(c) shows the SAR processor result, based on the Matched Filter. The amplitude grows from blue to red. The Fourier based approach was not able to separate the seven corner reflectors, anyhow it could separate two targets in azimuth which agrees to the fact that the azimuth resolution is higher than range resolution. Fig.4(d) shows the output of the proposed post-processing chain based on MUSIC when the selected data of Fig.4(c) is given as input. In this case all the seven corners reflector could be separated. As a validation of a correct separation a comparison between the detected geometry and the ground measurements has been carried out. More precisely, the average *estimated* slant range separation is $\pm 1.2m$, the *measured* slant range separation is $\pm 1.0m$ (the ground range measurement was projected into the slant range using the geometry of the acquisition), the average *estimated* azimuth separation is $\pm 0.86m$ and the *measured* azimuth separation is $\pm 0.80m$. The comparison results indicate that the estimation was reasonably successfull. As expected, the amplitude of the response of the corners can not be used as an estimation of the backscattered energy.

V. Conclusion

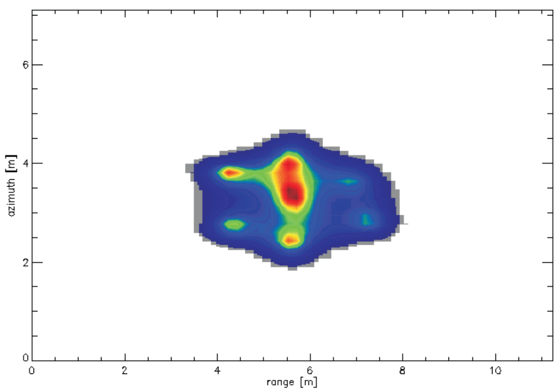
A post-processing approach for 2D SAR images based on the MUSIC estimator has been developed in order to separate



(a) SAR image. (b) Ground geometry.



(c) Focused image with Matched Filter.



(d) MUSIC focusing.

Fig. 4. Real data experiment.

targets within the Fourier resolution. Simulation showed good performance for separations lower than half the Fourier resolution. Concerning experimental results, it has been shown that all the targets have been detected with the correct geometry. In this frame, further investigation on test sites where strong backscatterers are present, such as urban areas, are suggested.

The limitation on the number of acquisitions ($K < \infty$) have implication in other practical conditions necessary for MUSIC to work. With the ideal infinite acquisition, the true covariance matrix will have exactly $N = L - D$ lower eigenvalues equal to the noise variance, independently of the SNR specified. In this way the number of targets D is known and the matrix $\mathbf{E}_N$ can be constructed. With the finite acquisition, the $L - D$ lower eigenvalues will differ from each other in a way that an external estimation of the number of targets is necessary. The SNR will make an important difference, since it will be related to the distance between the D higher eigenvalues to the $L-D$ lower ones. In [7] is given a detail statistical analysis of the MUSIC estimator in terms of the Cramer-Rao Bound. It is shown that the variance of MUSIC gets closer to the Cramer-Rao bound when $K \to \infty$ *and* $L \to \infty$, so even with an ideal acquisition, the spacial sampling L of the data must be large enough, when compared with the number of targets. In this sense, the spatial smoothing method presents a trade-off point since, in this case, $K = (M - m_1 + 1)(N - m_2 + 1)$ and $L = m_1 m_2$ so K and L are unfortunately inversely related. This means that for a given number of targets, the submatrix size must be large enough to achieve high value of L (and consequently a low variance which is a measure of resolution) but small enough to generate large K values (a good estimation of the covariance matrix). In the simulation and real data results the size of the subarray was chosen to agree with these conditions. Concerning the SNR, one important constraint observed regards the bandwidth of the chosen chirp in comparison with the image data. Another important constraint in the processing is the observation of the definition of the deramping signal, in order to cover all the image in a way that, its bandwidth does not reach the half of the sampling rate (aliasing).

In this work, an a priori knowledge of the number of sources has been exploited. Anyhow, external estimation methods [7] can be considered in order to avoid this a priori information knowledge. Estimation of the true reflectivity and other forms of estimation of the covariance matrix rather than (3) that gives more acquisitions such as the Forward-Backward method [1], [3], [6] are subject of further work.

VI. Acknowledgements

The authors would like to thank the project CAPTAER and the foundation FCMF, Brazil, for financing the travel of Philip Thompson to Germany.

References

[1] Stuart R. DeGraaf. *SAR Imaging via Modern 2-D Spectral Estimation Methods*. *IEEE Transactions on Image Processing*, 7,NO. 5:729–761, 1998.

[2] Marc Bara Iniesta. Appendix *A super-resolution method in airbone SAR* of the doctor thesis *Airbone SAR Interferometric Techniques for mapping Applicattions*. *Departament del Teoria del Senyal i Cominicacions, Universitat Poltcnica de Catalunya*, 2000.

[3] E. Barnard J. W. Odendaal and C. W. I. Pistorius. *Two Dimensional Superresolution Radar Imaging Using the MUSIC algorithm*. *IEEE Transactions on Antennas and Propagation*, 42,NO. 10:1386–1391, 1994.

[4] L. Marsili S. Barbarossa and G. Mungari. *SAR Super-Resolution Imaging by Signal Subspace Projection Techniques*. *Proc. of EUSAR'96*, pages 267–270, 1996.

[5] Ralph O. Schmidt. *Multiple Emitter Location and Signal Parameter Estimation*. *IEEE Transactions on Antennas and Propagation*, AP-34,NO. 3:276–280, 1986.

[6] Petre Stoica and Randolph Moses. *Introduction to Spectral Analysis*. Prentice-Hall, New Jersey, 1997.

[7] Petre Stoica and Arye Nehorai. *MUSIC, Maximum Likelihood, and Cramer-Rao Bound*. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37,NO. 5:720–741, 1989.