DTU Library

Restoration of polarimetric SAR images using simulated annealing

Schou, Jesper; Skriver, Henning

[Link back to DTU Orbit](#)

Restoration of Polarimetric SAR Images Using Simulated Annealing

Jesper Schou and Henning Skriver

Abstract—Filtering synthethic aperture radar (SAR) images ideally results in better estimates of the parameters character- izing the distributed targets in the images while preserving the structures of the nondistributed targets. However, these objectives are normally conflicting, often leading to a filtering approach favoring one of the objectives. An algorithm for estimating the radar cross-section (RCS) for intensity SAR images has previously been proposed in the literature based on Markov random fields and the stochastic optimization method simulated annealing. A new version of the algorithm is presented applicable to multilook polarimetric SAR images, resulting in an estimate of the mean co- variance matrix rather than the RCS. Small windows are applied in the filtering, and due to the iterative nature of the approach, reasonable estimates of the polarimetric quantities characterizing the distributed targets are obtained while at the same time preserving most of the structures in the image. The algorithm is evaluated using multilook polarimetric L-band data from the Danish airborne EMISAR system, and the impact of the algorithm on the unsupervised $H-\alpha$ classification is demonstrated.

Index Terms—Complex Wishart distribution, maximum likeli- hood estimation (MLE), polarimetry, simulated annealing, speckle filtering, synthetic aperture radar (SAR), unsupervised classifica- tion.

I. INTRODUCTION

A FULLY polarimetric SAR system measures the elements of the scattering matrix for each resolution cell. Assuming a homogeneous surface, the data are completely described by the complex covariance matrix (CM), which carries the full po- larimetric information [1]. Data are often multilook processed by averaging neighboring CMs to reduce the speckle present in SAR images. Standard multilook processing provides better estimates of the polarimetric quantities, thus facilitating subse- quent image analysis tasks, e.g., classification, segmentation, and visual inspection. As the structures in the image need to be maintained, there is an upper bound on the degree of multi- looking performed, and often specialized speckle reduction fil- ters are applied subsequently.

Several papers have described filtering of polarimetric data. The filtering techniques can result in a single intensity image, as for the polarimetric Whitening Filter by Novak and Burl [2] and the Multilook Polarimetric whitening filter described by Lopès and Séry [3] and Liu *et al.* [4], or in speckle reduced images of the co- and cross-polarized channels, as in the work by Lee *et al.* [5]. Also, the filtering techniques can preserve the entire polarimetric information by estimating the CM. This is done by Lee *et al.* [6] and by Lopès and Séry [3], where in the latter case, textural information is also preserved. Oliver and Quegan [7] derive a minimum mean square error (MSE) reconstruction filter capable of reconstructing the CM in the case of texture. The normal filtering approach is to apply a moving $M \times M$ window (although edge-directed windows have also been ap- plied [6]) and base the result upon the values inside the window. The size of the applied window depends on the number of looks of the input data, the desired variance of the estimates, and the need to maintain structures in the image.

In this paper, we present a new polarimetric filtering algo- rithm, where the mean CM is estimated. The algorithm is based on the filters developed by White [8], [9], McConnell *et al.* [10], McConnell and Oliver [11], and Oliver and Quegan [7], where the radar cross-section (RCS) is estimated from single-channel intensity data. By using Markov random fields (MRF) and sto- chastic relaxation algorithms, very high speckle reduction is achieved, while most of the structures in the image are pre- served.

The algorithm described in this paper belongs to the class of image restoration algorithms, where the goal is to restore the true process from an observed process. Using Bayesian methods, the image is modeled as a random field, and the image restoration problem can be expressed as an estimation problem. As the individual pixel values generally depend on the surrounding pixels, an MRF is chosen due to its ability to model spatial dependence [12]. By using the fact that the MRF follows a Gibbs distribution, the estimation problem can be expressed as an energy minimization problem, which is solved using the stochastic relaxation algorithm called simulated annealing (SA) [12], [13].

One of the most important applications of polarimetric SAR images is classification of land use areas, and several supervised and unsupervised classification schemes have been proposed in the literature [14], [15], [6], [16]. In this work, the polarimetric restoration is shown to be an efficient pre-processing step for the unsupervised classification scheme suggested by Cloude and Pottier [15].

The paper is organized as follows. Section II describes the concept of random field modeling of images, together with a general description of SA. In Section III, the complex Wishart distribution is used to derive the maximum likelihood estimate (MLE) of the mean CM, an estimate that is used in an optimiza- tion process to restore the mean CM. The restoration process is
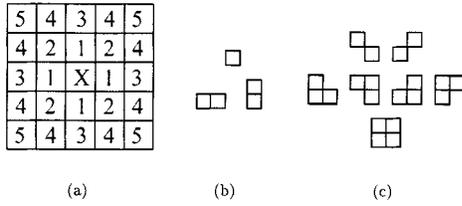
Fig. 1. (a) The n-order neighborhood of the central site (denoted by "x") contains the sites with numbers less than or equal to n, (b) cliques for a first-order neighborhood, and (c) additional cliques for a second-order neighborhood [13].

evaluated using multilook polarimetric SAR data from an agricultural site in Section IV. In Section V, an unsupervised classifier is applied to the restored polarimetric data set, and Section VI contains the conclusions.

II. RANDOM FIELDS AND ENERGY MINIMIZATION

We first briefly describe the Markov random field modeling of images and the use of Bayesian estimation methods. Then the energy minimization algorithm simulated annealing is discussed. The description in this section follows Li [13].

A. Markov Random Fields

Let $\mathcal{S} = \{s_1, s_2, \ldots, s_N\}$ denote the N sites of a 2-D image. The pixel values $\mathbf{x} = \{x_1, \ldots, x_N\}$ are realizations of the stochastic variables $\mathbf{X} = \{X_1, \ldots, X_N\}$ and the set of variables $\mathbf{X}$ defined on $\mathcal{S}$ is called a random field. $P(\mathbf{X} = \mathbf{x})$ is the probability of a particular configuration (image) of $\mathbf{X}$ [13], and all the possible configurations are called Ω.

The set of sites neighboring i, $\mathcal{N}_i$ is a subset of $\mathcal{S}$ having the following properties [13]:

$$i \notin \mathcal{N}_i \tag{1}$$

$$i' \in \mathcal{N}_i \Leftrightarrow i \in \mathcal{N}_{i'} \tag{2}$$

where the first condition states, that a site is not a neighbor to itself, and the second that the neighboring relationship is mutual. The collection of all the neighbor sets is called the neighborhood system $\mathcal{N} = \{\mathcal{N}_i | \forall\, i \in \mathcal{S}\}$, and the order of the neighborhood system can be seen in Fig. 1(a). A clique is defined as a subset of $\mathcal{S}$, for which every pair of sites are neighbors, and the set of cliques corresponding to the first-order and second-order neighborhoods are shown in Fig. 1(b) and (c), respectively.

Using these definitions we have that the random field $\mathbf{X}$ defined on $\mathcal{S}$ is a Markov random field (MRF), and using the Markov–Gibbs equivalence, the joint distribution of the MRF follows a Gibbs distribution [12], [13]

$$P(\mathbf{X} = \mathbf{x}) = \frac{1}{Z}\, e^{-[U(\mathbf{x})/T]}, \qquad \mathbf{x} \in \Omega \tag{3}$$

where T is called the temperature, and the normalization function Z, denoted the partition function, is defined as

$$Z = \sum_{\mathbf{x} \in \Omega} e^{-[U(\mathbf{x})/T]}. \tag{4}$$

The energy function $U(\mathbf{x})$ depends on the pixel values contained in the cliques, $U(\mathbf{x})$ is highly dependent on the specific image

analysis task, and for mean CM estimation, it is described in Section III.

Bayesian estimation methods are used intensively in image analysis for retrieving information on a process underlying an observed process. If the underlying process is called $\mathbf{X}$ and the observed process $\mathbf{Y}$, then Bayes rule states that

$$P(\mathbf{x}|\mathbf{y}) \propto P(\mathbf{y}|\mathbf{x})P(\mathbf{x}) \tag{5}$$

with $P(\mathbf{y}|\mathbf{x})$ being the likelihood function, $P(\mathbf{x})$ the *a priori* probability, and $P(\mathbf{x}|\mathbf{y})$ the *a posteriori* conditional probability. Modeling $\mathbf{X}$ as a MRF following a Gibbs distribution, we can rewrite the *a posteriori* probability as

$$P(\mathbf{x}|\mathbf{y}) = \frac{1}{Z}\, e^{-[U(\mathbf{x}|\mathbf{y})/T]} \tag{6}$$

using the *a posteriori* conditional energy function $U(\mathbf{x}|\mathbf{y})$. The configuration $\hat{\mathbf{x}}$ that maximizes the *a posteriori* conditional probability is called the maximum *a posteriori* (MAP) estimate. If knowledge about the *a priori* probability is omitted, the resulting estimate is the MLE. For the mean CM restoration algorithm in this paper, the aim is to estimate the underlying, unspeckled image $\mathbf{X}$ from the observed, speckled image $\mathbf{Y}$ using the MLE, as discussed in Section III.

B. Energy Minimization Using Simulated Annealing

Finding the global maximum of the *a posteriori* distribution in (6), or equivalently the global minimum of the energy function is often made difficult by local maxima (minima), which can easily trap the optimization algorithm. In the present work, the stochastic optimization method SA is applied to find the global minimum of the energy function.

SA searches for the configuration $\mathbf{x}$ that minimizes an (multidimensional) energy function $U(\mathbf{x})$. It is iterative by nature, where a new configuration $\mathbf{x}^j$ for iteration j is found from the previous configuration $\mathbf{x}^{j-1}$ by applying a generation mechanism and accepting the new configuration using an acceptance criterion based on the energy divergence $\Delta U = U(\mathbf{x}^j) - U(\mathbf{x}^{j-1})$. The temperature T controls the optimization, and it is decreased throughout the optimization process. For the first iterations when T is high, there is a high probability of accepting configurations resulting in an increase in the energy, thus making SA able to get out of local minima. As the temperature is gradually decreased, the probability of accepting configurations resulting in increasing energies is reduced, and at the end of the minimization no increases are accepted, and the global minimum configuration is ideally reached.

The SA algorithm can be described as in [13], [17].

1) Initialize T and $\mathbf{x}^0$.
2) Iteration j
    a) Propose updated configuration $\mathbf{x}^j$.
    b) Compute $U(\mathbf{x}^j)$.
    c) Accept/reject $\mathbf{x}^j$ based on ΔU and T.
    d) Stop if stop criterion is fulfilled.
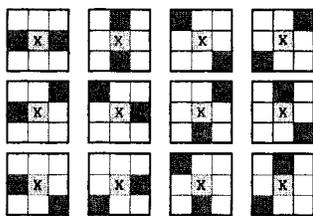   e) Decrease T according to cooling schedule.
    f) Next iteration.

Fig. 2. Set of 12 connection filters used by the generation mechanism to find the new estimate of **C**. "x" denotes the central pixel, and each filter has two neighbors to the central pixel.

Since the elements in the covariance matrix are continuous rather than belonging to a discrete set, a continuous minimization method is needed to estimate **C**, and this is reflected in the choice of generation mechanism in step 2 a). In the original work by White [8], [9], the new pixel value, which in White's case reflected the RCS, is derived from a set of 12 connection filters. Thus, instead of choosing from an infinite number of pixel values, the generation mechanism only needs to choose between 12 different values. The set of connection filters, which is a subset of the possible cliques for a fifth-order neighborhood system, is seen in Fig. 2. The central pixel has two neighbors for each of the 12 filters, and for each filter the new value of the central pixel is based on the original value of the central pixel and the present values of the neighboring pixels. In the following, a more general description of the filters is proposed, where the number of connection filters is denoted N_c and the number of neighbor pixels within each filter is denoted n_n, i.e., $N_c = 12$ and $n_n = 2$ for Fig. 2.

The Metropolis criterion is applied as the acceptance criterion, where a new configuration resulting in a negative energy divergence is always accepted, while a configuration resulting in positive divergence is accepted with probability $e^{-\Delta U/T}$ [13]. As the stop criterion, Hellwich [17] suggests using either a given number of iterations or when an insignificant number of changes is made over a given number of iterations. SA finds the global minimum of $U(\mathbf{x})$ provided that the temperature is cooled sufficiently slowly toward zero. The cooling schedule needed is, however, too slow to be of any practical use, and we apply the following logarithmic cooling schedule [7], [13], [18]

$$T_j = \frac{T_0}{\ln(1 + \beta j)} \tag{7}$$

with j being the iteration number, T_0 the starting temperature that should be "large enough" [19], and β a constant.

III. COVARIANCE MATRIX RESTORATION

In this section, we first describe the complex Wishart distribution, and then this distribution is used in the optimization process to restore the mean covariance matrix from the multi-looked polarimetric SAR data. These latter data will in the following be denoted as the original data. Finally, a bias in the mean CM estimate is discussed.

A. Complex Wishart Distribution

A polarimetric SAR transmits horizontally and vertically polarized signals alternately, and by measuring both polarizations

of the backscattered signal, the complex scattering matrix is formed. In the following, we will assume that reciprocity [1], [20] holds, i.e., $S_{hv} = S_{vh}$ (for backscattering alignment) with S_{pq} being the complex scattering amplitude for receive polarization p and transmit polarization q, thus the single-look polarimetric data at each pixel can be represented using the complex scattering vector **s**

$$\mathbf{s} = (S_{hh} \quad S_{hv} \quad S_{vv})^t \equiv (S_1 \quad S_2 \quad S_3)^t \tag{8}$$

where t denotes transpose. The scattering vector from a homogeneous surface is a zero mean multivariate complex Gaussian random variable [21] and from **s**, the Hermitian sample CM $\boldsymbol{\Sigma}$ is formed as

$$\boldsymbol{\Sigma} = 4\pi \langle \mathbf{s}\, \mathbf{s}^{t*} \rangle = 4\pi \left\langle \begin{bmatrix} |S_1|^2 & S_1 S_2^* & S_1 S_3^* \\ S_2 S_1^* & |S_2|^2 & S_2 S_3^* \\ S_3 S_1^* & S_3 S_2^* & |S_3|^2 \end{bmatrix} \right\rangle \tag{9}$$

where $*$ denotes complex conjugation. The sample covariance matrix follows a complex Wishart distribution, which is a function of the number of looks L and the Hermitian mean CM **C** [3], [21]

$$p(\boldsymbol{\Sigma}|\mathbf{C}) = \frac{L^{Lp}|\boldsymbol{\Sigma}|^{L-p}e^{-L\,\mathrm{tr}(\mathbf{C}^{-1}\boldsymbol{\Sigma})}}{f(L,\,p)|\mathbf{C}|^L} \tag{10}$$

where p is the dimension of **s**, i.e., $p = 3$ assuming reciprocity, $|\cdots|$ and $\mathrm{tr}(\cdots)$ denote the determinant and trace, respectively, and $f(L,\,p)$ is given by

$$f(L,\,p) = \pi^{(1/2)p(p-1)} \prod_{j=1}^{p} \Gamma(L - j + 1) \tag{11}$$

with $\Gamma(\cdots)$ being the Gamma function. As shown by Goodman [21], $\boldsymbol{\Sigma}$ is an unbiased estimate of the mean CM. The mean CM **C** contains the full polarimetric information for a homogeneous surface, and represents the quantity we want to estimate in the CM restoration, and (10) thus represents the likelihood function, as described in Section II. It should be noted that multi-frequency, and/or multitemporal polarimetric data can also be described using the complex Wishart distribution, and hence, the remainder of this work is applicable to these kinds of polarimetric SAR data as well.

We assume homogeneous areas in this work. If texture is present, a possible approach would be to apply the product model, where the mean CM is the product of a texture variable and a mean CM from a homogeneous surface [3], [7], [22]. Lopés and Séry [3] derive various estimates of the texture variable, which must be included in the likelihood function.

B. ML and MAP Estimates of the Mean CM

Using the connection filters shown in Fig. 2, the new estimate of the mean CM for the central pixel is based on the original sample CM at the central pixel and the present estimates of the mean CM for the neighbors. Denoting the original, central

sample CM as $\mathbf{\Sigma}_0$ and the neighboring mean CM estimates as $\mathbf{Z}_1, \ldots, \mathbf{Z}_{n_n}$, we find the *a posteriori* probability as

$$
\begin{aligned}
P(\mathbf{C}_{ik}&|\mathbf{\Sigma}_0, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_n k})\\
&\propto P(\mathbf{\Sigma}_0, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_n k}|\mathbf{C}_{ik})P(\mathbf{C}_{ik})\\
&= [P(\mathbf{\Sigma}_0|\mathbf{C}_{ik})P(\mathbf{Z}_{1k}|\mathbf{C}_{ik})\cdots P(\mathbf{Z}_{n_n k}|\mathbf{C}_{ik})]P(\mathbf{C}_{ik}) \quad (12)
\end{aligned}
$$

where the suffix i denotes the site and the suffix k denotes the connection filter. The brackets encompasses the conditional probability density function, where independence between the neighboring pixels has been assumed. This assumption facilitates the derivation of the mean CM estimate considerably, but it has limited validity because on the one hand, the neighboring mean CM estimates $\mathbf{Z}_1, \ldots, \mathbf{Z}_{n_n}$ are initialized using the original sample CM, which correlates each of them with $\mathbf{\Sigma}_0$ assuming correlation between neighboring pixels in the original data, and on the other hand, the neighboring estimates become increasingly dependent as the algorithm iterates. We will discuss the effect of this assumption further after having derived the mean CM estimate. The original sample CM, which is included in the conditional energy function to ensure that the present estimate is consistent with the original data [7], and the neighboring mean CM estimates are all assumed to be realizations of a complex Wishart distribution having mean CM $\mathbf{C}_{ik}$. L is the number of looks for $\mathbf{\Sigma}_0$, and L is constant throughout the algorithm. The number of looks for the neighboring mean CM estimates, κ is estimated for each new pixel visited, as the estimates are continuously being updated, and κ is taken as the average of the locally estimated number of looks for each of the three polarizations, where the estimation is performed in a $M \times M$ window around the central pixel

$$
\hat{\kappa} = \frac{1}{3}\sum_j \frac{E\{\hat{\sigma}_j\}^2}{V\{\hat{\sigma}_j\}}, \qquad j = hh, \, hv, \, vv \quad (13)
$$

with $\hat{\sigma}_j$ being the backscattering coefficients for polarization j from the restored CM.

We first derive the MLE of the mean CM using (10) and (12). For each of the connections, the likelihood function at site i is given by

$$
\begin{aligned}
P(\mathbf{\Sigma}_0&, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_n k}|\mathbf{C}_{ik})\\
&= \frac{L^{Lp}|\mathbf{\Sigma}_0|^{L-p}\, e^{-L\mathrm{tr}\,(\mathbf{C}_{ik}^{-1}\mathbf{\Sigma}_0)}}{f(L, p)|\mathbf{C}_{ik}|^L}\\
&\quad \times \prod_{j=1}^{n_n} \frac{\kappa^{\kappa p}|\mathbf{Z}_{jk}|^{\kappa-p}\, e^{-\kappa\mathrm{tr}\,(\mathbf{C}_{ik}^{-1}\mathbf{Z}_{jk})}}{f(\kappa, p)|\mathbf{C}_{ik}|^{\kappa}}. \quad (14)
\end{aligned}
$$

By omitting terms not involving $\mathbf{C}_{ik}$ the log likelihood function is given as

$$
\begin{aligned}
Lf(\mathbf{C}_{ik}) &= \log P(\mathbf{\Sigma}_0, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_n k}|\mathbf{C}_{ik})\\
&= -\Big(L\mathrm{tr}\,\big(\mathbf{C}_{ik}^{-1}\mathbf{\Sigma}_0\big) + \kappa\,\big(\mathrm{tr}\big(\mathbf{C}_{ik}^{-1}\mathbf{Z}_{1k}\big) + \cdots\\
&\quad + \mathrm{tr}\big(\mathbf{C}_{ik}^{-1}\mathbf{Z}_{n_n k}\big)\big) + (L + n_n\kappa)\log|\mathbf{C}_{ik}|\Big). \quad (15)
\end{aligned}
$$

We find the MLE of the mean CM by solving $dLf(\mathbf{C}_{ik})/d\mathbf{C}_{ik} = \mathbf{0}$, with $\mathbf{o}$ being the 3×3 null

matrix. Using expressions for the derivative of trace $d\mathrm{tr}\,(\mathbf{X}^{-1}\mathbf{Y})/d\mathbf{X} = -\mathbf{X}^{-t}\mathbf{Y}^t\mathbf{X}^{-t}$, with $-t$ denoting inverse transpose, and the derivative of the logarithm of the determinant, $d\log(|\mathbf{X}|)/d\mathbf{X} = \mathbf{X}^{-t}$ [23], the MLE is found from

$$
L\mathbf{\Sigma}_0 + \kappa\mathbf{Z}_{1k} + \cdots + \kappa\mathbf{Z}_{n_n k} - (L + n_n\kappa)\hat{\mathbf{C}}_{ik} = \mathbf{0} \quad (16)
$$

resulting in

$$
\hat{\mathbf{C}}_{ik} = \frac{\mathbf{\Sigma}_0 + \dfrac{\kappa}{L}(\mathbf{Z}_{1k} + \cdots + \mathbf{Z}_{n_n k})}{1 + n_n\dfrac{\kappa}{L}}. \quad (17)
$$

From this equation it can be seen, that the locally estimated number of looks, κ, controls the degree of averaging. For a smooth surface κ is high, resulting in $\hat{\mathbf{C}}_{ik} \approx (1/n_n)(\mathbf{Z}_{1k} + \cdots + \mathbf{Z}_{n_n k})$, whereas a discontinuity in the local $M \times M$ window makes κ small, and thus gives a higher priority to the original sample CM. This expression is a generalization of the result given by McConnell *et al.* [10], McConnell and Oliver [11], and Oliver and Quegan [7], where they find the MLE for the RCS, using $n_n = 2$, as

$$
\hat{\sigma}_{ik} = \frac{I_0 + \dfrac{\kappa}{L}(x_{1k} + x_{2k})}{1 + 2\dfrac{\kappa}{L}} \quad (18)
$$

with $\hat{\sigma}_{ik}$ being the estimated RCS, I_0 the original sample intensity, and x_1 and x_2 the current estimates of the RCS at the neighboring pixels. Estimating the RCS can be seen as a special case of mean CM estimation, as the RCS are the diagonal terms in the covariance matrix. As previously mentioned, the estimate $\hat{\mathbf{C}}_{ik}$ assumes that the neighboring CMs are uncorrelated. One approach to correct this assumption could be, to include a factor describing the effective number of independent samples represented by the neighboring sites; this, however, has not been attempted in this work.

For obtaining the MAP estimate of the mean CM, the *a priori* probability, $P(\mathbf{C})$, needs to be included. The problem is, that this distribution is not known, making it impossible to derive the MAP estimate. It is however noticeable, that if $\mathbf{C}$, for lack of a known distribution, is assumed to follow a complex Wishart distribution, having ζ number of looks and the "true" mean CM $\mathbf{M}$, then the log likelihood function for the MAP estimate using $n_n = 2$ is given as

$$
\begin{aligned}
Lf(\mathbf{C}_{ik}) &= \log P(\mathbf{\Sigma}_0, \mathbf{Z}_{1k}, \mathbf{Z}_{2k}|\mathbf{C}_{ik}) + \log P(\mathbf{C}_{ik})\\
&= -\big(L\mathrm{tr}\,\big(\mathbf{C}_{ik}^{-1}\mathbf{\Sigma}_0\big) + \kappa\mathrm{tr}\,\big(\mathbf{C}_{ik}^{-1}\mathbf{Z}_{1k}\big) + \kappa\mathrm{tr}\,\big(\mathbf{C}_{ik}^{-1}\mathbf{Z}_{2k}\big)\\
&\quad + \zeta\mathrm{tr}\,\big(\mathbf{M}^{-1}\mathbf{C}_{ik}\big) + (L + 2\kappa + p - \zeta)\log|\mathbf{C}_{ik}|\big). \quad (19)
\end{aligned}
$$

Using the following expression $d\mathrm{tr}(\mathbf{Y}^{-1}\mathbf{X})/d\mathbf{X} = \mathbf{Y}^{-t}$ [23] and the ones mentioned previously, we obtain the MAP estimate of $\mathbf{C}_{ik}$ as the solution for $n_n = 0$

$$
\begin{aligned}
\zeta\hat{\mathbf{C}}_{ik}\mathbf{M}^{-1}\hat{\mathbf{C}}_{ik} &+ \hat{\mathbf{C}}_{ik}(2\kappa + L + p - \zeta) - L\mathbf{\Sigma}_0\\
&- \kappa(\mathbf{Z}_{1k} + \mathbf{Z}_{2k}) = \mathbf{0}. \quad (20)
\end{aligned}
$$

This quadratic matrix equation is, again, a generalization of the MAP estimate of the RCS [7], which is given by the solution to the scalar quadratic expression for $n_n = 0$

$$\nu\hat{\sigma}_{ik}\mu^{-1}\hat{\sigma}_{ik}+\hat{\sigma}_{ik}(2\kappa+L+1-\nu)-LI_0-\kappa(x_1+x_2) = 0 \quad (21)$$

where μ and ν are the mean RCS and order parameter, respectively, and the rest of the parameters are similar to equation (18). The MAP estimate for the mean CM is not used in the following, because, on the one hand, the assumption of $\mathbf{C}$ following a complex Wishart distribution is not supported, either in this work or in the literature, and, on the other hand, the solution to the quadratic matrix equation (20) is far from straightforward [24].

C. Mean CM Restoration Using Simulated Annealing

The mean CM restoration is performed as a Gibbs energy minimization process, with the likelihood and *a posteriori* energy functions given as

$$U(\boldsymbol{\Sigma}_0, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_nk}|\mathbf{C}_{ik})$$
$$= -\ln(P(\boldsymbol{\Sigma}_0, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_nk}|\mathbf{C}_{ik})) \quad (22)$$
$$U_{ML}(\mathbf{C}_{ik}|\boldsymbol{\Sigma}_0, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_nk})$$
$$= U(\boldsymbol{\Sigma}_0, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_nk}|\mathbf{C}_{ik}) \quad (23)$$

From (22) and (23) it is verified, that a high likelihood probability results in a small *a posteriori* energy; a necessary condition when the restoration is performed as an energy minimization process. For each of the connections, we calculate the probability

$$P\left(\hat{\mathbf{C}}_{ik}|\boldsymbol{\Sigma}_0, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_nk}\right)$$
$$= \frac{1}{Z}e^{-U_{ML}(\hat{\mathbf{C}}_{ik}|\boldsymbol{\Sigma}_0, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_nk})/T} \quad (24)$$

with the partition function given as

$$Z = \sum_k e^{-U_{ML}(\hat{\mathbf{C}}_{ik}|\boldsymbol{\Sigma}_0, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_nk})/T} \quad (25)$$

and identify the connection k_0 resulting in the highest *a posteriori* probability (the lowest *a posteriori* energy). This connection is then accepted or rejected based on the Metropolis criterion, as previously mentioned in Section II-B.

D. Bias of the Mean CM Estimate

In the case of RCS reconstruction, Oliver and Quegan [7] notice a bias in the estimate of the RCS, resulting in lower values of $\hat{\sigma}$ than expected, and we observe a bias for the restored CM as well. The bias of the estimate is a result of the choice made by the algorithm between the different connection filters. As discussed in Section III-C, the connection filter resulting in the highest *a posteriori* probability is chosen from the set of connection filters. If the mean CM estimate is unbiased, the average of the n_n neighboring covariance matrices within the connection filter must be equal to the average of the original sample CM, as this in turn is an unbiased estimate of the mean CM [21]. We will

show, however, that the connection filter resulting in maximum *a posteriori* probability leads to a biased mean CM estimate.

From Section IV-C we have that connection filter k_0 resulting in MLE $\hat{\mathbf{C}}_{k_0}$ is chosen from the set of N_c possible connection filters iff

$$P\left(\boldsymbol{\Sigma}_0, \mathbf{Z}_{1k_0}, \ldots, \mathbf{Z}_{n_nk_0}|\hat{\mathbf{C}}_{k_0}\right)$$
$$\geq P\left(\boldsymbol{\Sigma}_0, \mathbf{Z}_{1k}, \ldots, \mathbf{Z}_{n_nk}|\hat{\mathbf{C}}_k\right), \qquad k_0, k \leq N_c, k_0 \neq k. \quad (26)$$

The overall most likely connection filter is the filter where the n_n neighboring CMs result in the overall highest probability (26), and these values are found by solving the log likelihood equation $[d\log P(\boldsymbol{\Sigma}_0, \mathbf{Z}_{1k_0}, \ldots, \mathbf{Z}_{n_nk_0}|\hat{\mathbf{C}}_{k_0})]/d\mathbf{Z}_{1k_0}\cdots d\mathbf{Z}_{n_nk_0} = \mathbf{0}$. To be able to solve this equation, we need to introduce a constraint. In Section IV-E it is shown that the backscattering coefficients of the restored CM are biased with approximately the same amount, whereas the complex correlation coefficients are practically unbiased, hence the bias must be of the form $\langle\hat{\mathbf{C}}\rangle = constant \times \langle\boldsymbol{\Sigma}_0\rangle$. We therefore define the neighboring CM's by means of the original central sample CM, i.e., $\mathbf{Z}_{jk_0} = v_{jk_0}\boldsymbol{\Sigma}_0$, with the factors v_{jk_0} being real scalars. Solving the log likelihood equation with respect to these factors, i.e., $[d\log P(\boldsymbol{\Sigma}_0, \mathbf{Z}_{1k_0}, \ldots, \mathbf{Z}_{n_nk_0}|\hat{\mathbf{C}}_{k_0})]/dv_{1k_0}\cdots dv_{n_nk_0} = \mathbf{0}$, yields

$$v_{jk_0} = \frac{L(\kappa - p)}{\kappa(L + n_np)}, \qquad j \leq n_n \quad (27)$$

This shows, that the overall highest probability (26) is obtained, if the n_n neighbors within the connection filter are given as $\mathbf{Z}_{jk_0} = [L(\kappa - p)/\kappa(L + n_np)]\boldsymbol{\Sigma}_0$. The mean CM estimate from this specific filter is biased as $L(\kappa - p)/\kappa(L + n_np) < 1$, and to derive the expression of the bias resulting from this filter, we insert $\mathbf{Z}_{jk_0}$ in the MLE for the mean CM (17) giving

$$\langle\hat{\mathbf{C}}\rangle = \frac{L}{L + n_np}\langle\boldsymbol{\Sigma}_0\rangle. \quad (28)$$

This expression only depends on the number of looks of the original data, L, on the number of neighbors within the connection filter, n_n, and on the dimension of the CM, p. It should be mentioned that for the RCS reconstruction algorithm the log likelihood equation $dP(I_0, x_{1k_0}, \ldots, x_{n_nk_0}|\hat{\sigma}_{k_0})/dx_{1k_0}\cdots dx_{n_nk_0} = 0$ can be solved using the appropriate Gamma distributions and without introducing any constraint, resulting in $\langle\hat{\sigma}\rangle = L/(L+n_n)\langle I_0\rangle$, i.e., similar to (28) with $p = 1$.

IV. EXPERIMENTAL RESULTS

In this section the mean CM restoration algorithm is evaluated using airborne EMISAR data, with focus on the preservation of polarimetric information and structures. In addition to the restoration algorithm, the polarimetric SAR speckle filter by Lee *et al.* [6] is applied for comparison.
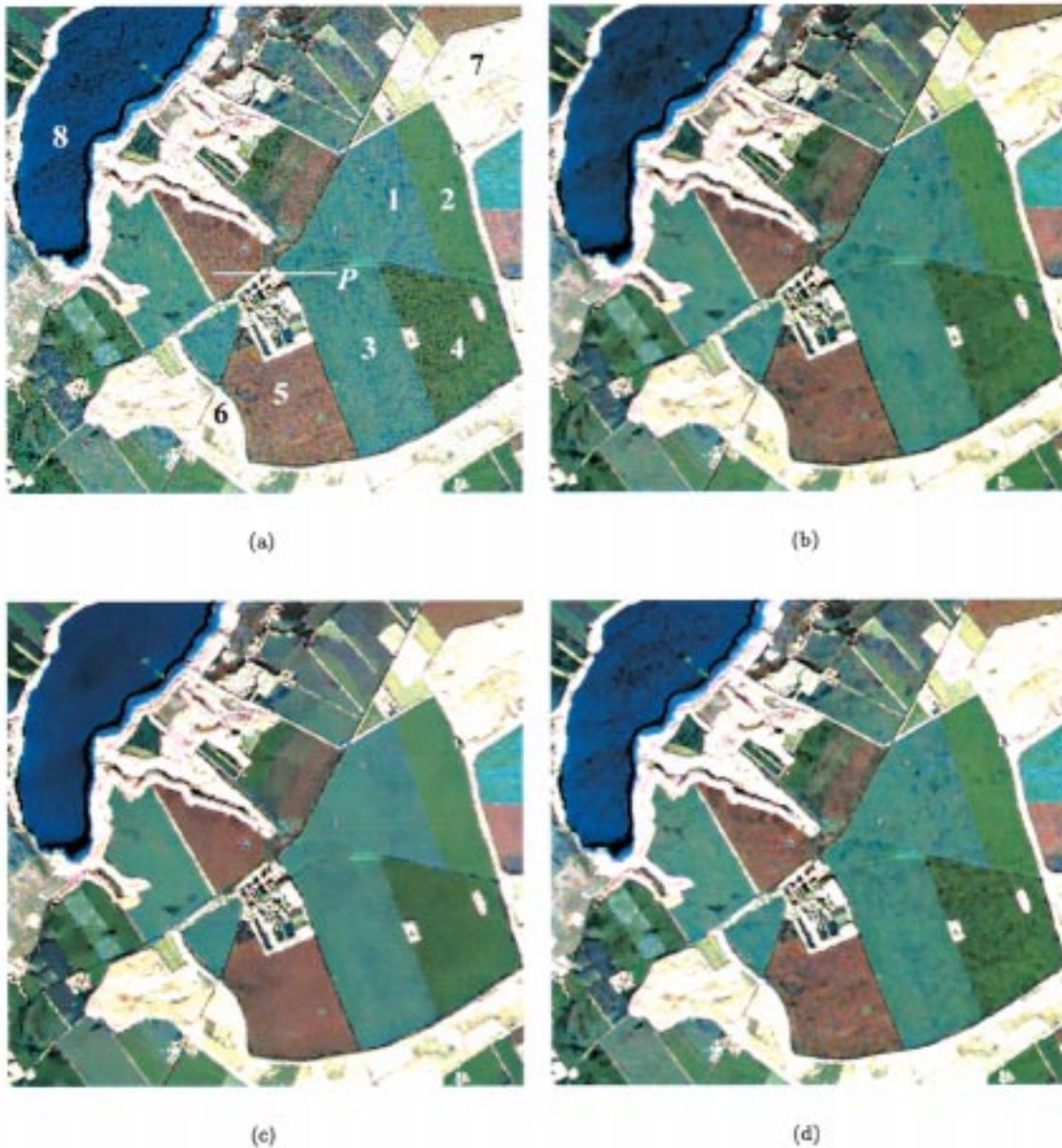
Fig. 3.   Covariance matrix estimation using 13-look polarimetric C-band EMISAR data, color coded using hv (red), hh (green), and vv (blue). (a) Original data including test areas and profiles, (b) mean CM restored data using $N_c = 12$ filters, (c) mean CM restored data using filter banks, each having $N_c = 4$, and (d) polarimetric SAR speckle filtered data.

A. SAR Data

The mean CM restoration algorithm is evaluated using fully polarimetric C-band data from the Danish airborne polarimetric SAR, EMISAR, developed at the Section of Electromagnetic Systems (EMI), Orsted, DTU [25]. EMISAR operates at L- and C-band (1.25 GHz and 5.3 GHz, respectively). Flown on board a Gulfstream G-3 aircraft from the Royal Danish Air Force at an altitude of typically 12.5 km, it acquires data with a ground range swath of 12 km and incidence angles ranging from $35°$ to $60°$. The test image covers approximately 8 km^2 (550×580 pixels), and includes several agricultural fields, forests, buildings, roads, and natural vegetation. The area is located in Jutland, Denmark, and a RGB image using the hv (red), hh (green), and vv (blue) polarizations is shown in Fig. 3(a) (this image, and

the following EMISAR images, are all shown in logarithmic scale). The test area is a subset of the area shown by Christensen *et al.* [25] and Skriver *et al.* [26], where RGB versions using the same color coding for both frequencies are shown in Figs. 8 and 1, respectively. In the following the EMISAR 13-look covariance data having a 5×5 m ground pixel spacing are used for the mean CM restoration algorithm. These data are denoted as the original data.

B. Estimated Polarimetric Parameters

The 12 connection filters from Fig. 2 are applied in Fig. 3(b), where an RGB image using $\hat{\sigma}_{hh}$, $\hat{\sigma}_{hv}$, and $\hat{\sigma}_{vv}$ resulting from the mean CM restoration is shown, using 100 iterations and a local

3×3 window for estimating κ (note that we restore all the elements of the mean CM, and hence, the diagonal elements $\hat{\sigma}_j$ are only part of the result). It is seen that the homogeneous areas in the restored image, especially the agricultural fields, contains a number of small patches having different intensity levels. The patches in the restored image, which do not disappear when applying more iterations, reflect small areas in the original image that appear either darker or brighter than their surroundings. Hence, they are caused by the restoration algorithm being highly sensitive to the original data. Although the patches in Fig. 3(b) reflects physical changes in the original data, a more homogeneous result is desirable for many applications. One method for obtaining this would be to include more neighbors within each filter, i.e., increase n_n, but this increases the already significant computational load of the algorithm. Another method is to apply fewer connection filters in the restoration, i.e., reduce N_c, as this reduces the capabilities of the algorithm to track and maintain the small patches throughout the iterations, but this also results in loss of some of the actual structures in the scene as fewer filter orientations are examined. By applying an alternate subset of the connection filters in the restoration, however, it is possible to obtain a high degree of filtering for the homogeneous areas, while maintaining the actual structures.

We divide the 12 connection filters in Fig. 2 in three filter banks, where each bank consists of the four filters in a row. For a specific iteration, we apply a given filter bank, and we alternate between the three banks for each new iteration. A structure oriented according to one of the connection filters in bank A results in a high *a posteriori* probability (24), where the partition function Z is derived from the $N_c = 4$ filters in the bank. The filters in the other filter banks are oriented differently and result in lower *a posteriori* probabilities, and are thus likely to be rejected based on the Metropolis criterion, hence a new estimate is only obtained when filter bank A is applied. For the homogeneous areas no preferred orientation exists, and each of the filter banks are likely to generate a new estimate of the mean CM.

The restoration estimate in Fig. 3(c) is produced using this procedure, and it is verified that the fields appear very smooth, while practically all of the actual structures have been preserved. It should be stressed that applying an alternate subset of connection filters is theoretically fully acceptable. The set of filters in Fig. 2 is actually a subset of all the possible connection filters having $n_n = 2$ itself. The number of filters in each filter bank must be equal, however, as the partition function (25) is formed as the sum of *a posteriori* probabilities within the bank. Also, the filters in each bank should have significantly different orientations, so a filter oriented along a structure results in high *a posteriori* probability. Finally, it is desirable that the computational load of the algorithm is reduced by a factor of 3, approximately, by applying the filter bank approach as compared to applying all 12 connection filters.

The polarimetric SAR speckle filter proposed by Lee *et al.* [6] has been implemented for comparison, and Fig. 3(d) shows the filtering result. The polarimetric SAR speckle filter clearly reduces speckle in the homogeneous areas, although not to the same degree as the mean CM restoration filter. Also, the structures having small spatial dimensions seems a bit more blurred than the results using iterative filtering. However, quite a few

of the structures are maintained in the speckle filtered result, and especially the boundaries between different land-use classes (fields, forest, and natural vegetation) are preserved very well.

The mean CM restoration also provides an estimate of the complex correlation coefficient between channel i and j, $\rho_{i,j} = \langle S_i S_j^* \rangle / \sqrt{|S_i|^2 |S_j|^2}$, $i \neq j$. Fig. 4 shows the amplitude (coherence) and phase (phase difference) of $\rho_{hh,vv}$ for the original 13-look data and for the MLEs. The MLEs of the coherence has clearly less variance in the homogeneous areas than the original coherence. In general, the variance is higher for the estimates around the structures because the presence of a structure gives a higher priority to the original sample CM in the MLE, as discussed in Section III-B. Comparing the original phase difference image with the original coherence image, it is seen that the phase difference has the largest variance in areas with small coherence [7]. In Fig. 4(d) showing the estimated phase difference, we see that most of the fields appear very smooth, and the fields having low coherence appear somewhat inhomogeneous although a high degree of filtering has occurred.

C. Speckle Reduction

The estimated equivalent number of looks (ENL) for $\hat{\sigma}$ for a homogeneous area [field 1 in Fig. 3(a)] as the mean CM restoration algorithm iterates, is shown in Fig. 5. Also shown in Fig. 5 is the estimated number of looks for the noniterative polarimetric SAR speckle filter. The number of iterations used for a specific application is obviously a compromise between the desired degree of speckle reduction and the work load of the restoration. It should be noted, that an almost linear relation between the speed-ups and the number of processors is expected if applying parallel computing to the restoration algorithm [27].

D. Structure Preservation

When looking at the structures in Fig. 3(c), it is observed that the MLE preserves most of them very well. It is primarily the contrast of the structures to the surroundings and the dimensions of the structures, which affect the result. Structures having a large contrast are generally preserved, while small and/or narrow linear structures can be blurred, provided they have low contrast to the surroundings. The buildings in the image are examples of structures having large contrast to the surroundings, and they are all preserved without any visual blurring. Some of the roads, on the other hand, cause problems for the restoration process, partly due to the small spatial extent at the 5-m pixel spacing and partly because the surrounding fields have comparable intensity levels. Also, the orientation of the structures can influence the result. Structures having an orientation along one of the connection filters are more likely to survive the restoration than structures having other orientations, hence in general a symmetric set of filters should be applied.

Fig. 6 shows the profile [marked as P on Fig. 3(a)] for the original data, the polarimetric speckle filtered data, and the mean CM restored data. Obviously, the profiles support the statement that structures having a large contrast are preserved by the mean CM restoration algorithm, while structures having a lower contrast, such as the dirt road, can be blurred. The homogeneous fields all appear extremely smooth. In the profile
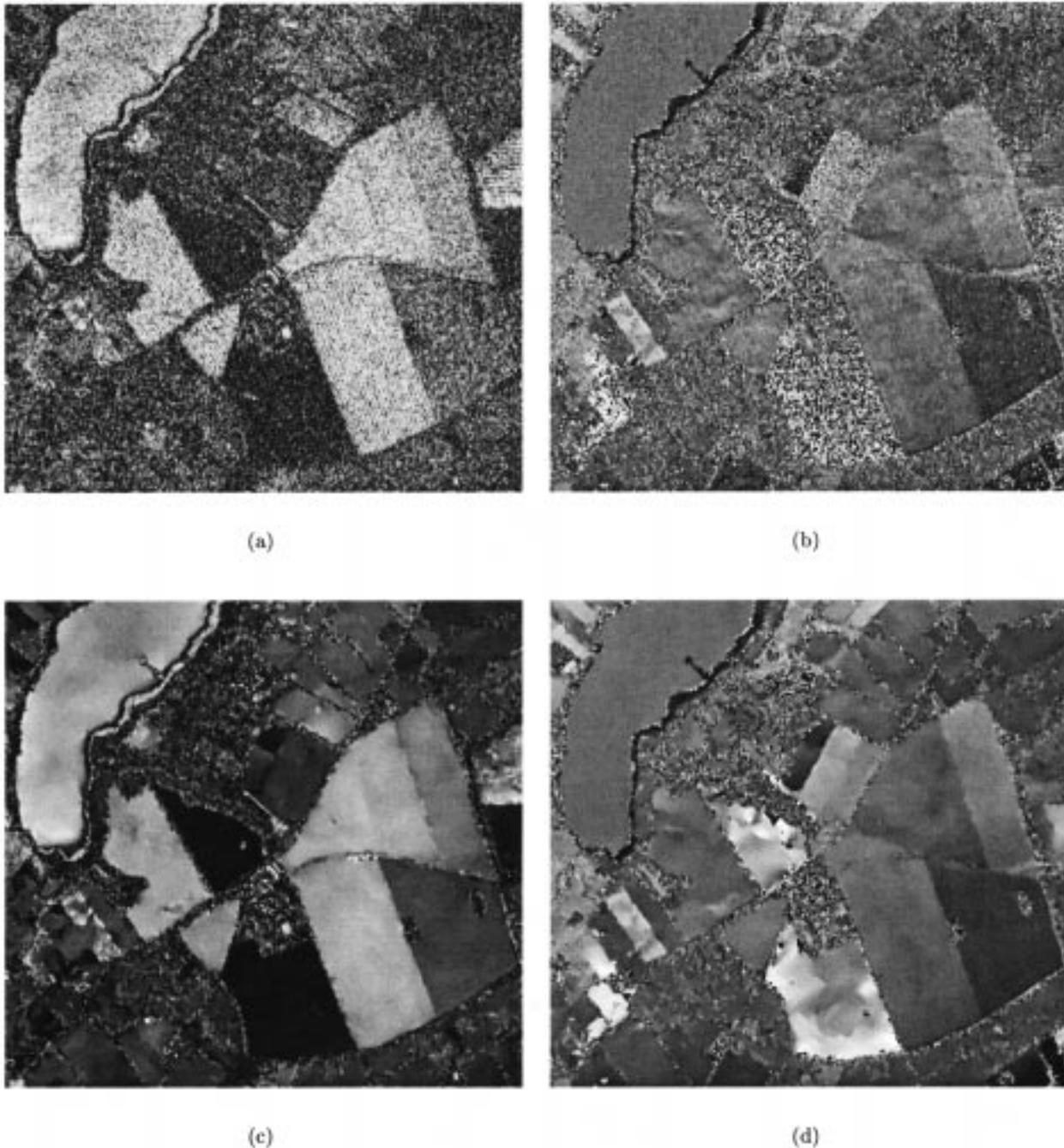
Fig. 4.   Mean covariance matrix estimation. (a) $|\rho_{hh,\,vv}|$ original, (b) $\angle\rho_{hh,\,vv}$ original, (c) $|\hat{\rho}_{hh,\,vv}|$, and (d) $\angle\hat{\rho}_{hh,\,vv}$.

from the polarimetric SAR speckle filter it is observed that not all of the valleys in the intensity profile for the building area are fully preserved, and that the fields appear less smooth. Finally, it should be stressed that structures appearing significant in one of the parameters of the CM may be less significant in the remaining parameters of the CM and it is the entire CM that has been applied by both filtering algorithms.

E. Accuracy of Estimated Parameters

To investigate the bias of the restored CM, we form the ratios between the corresponding backscattering coefficients in the restored and original data $r_j = \hat{\sigma}_j/\sigma_j$. The bias of the com-

plex correlation coefficient is examined by comparing the coherence and phase difference of the restored and original data, i.e., $\Delta_{|\rho_{hh,\,vv}|} = |\hat{\rho}_{hh,\,vv}| - |\rho_{hh,\,vv}|$ and $\Delta_{\angle\rho_{hh,\,vv}} = \angle\hat{\rho}_{hh,\,vv} - \angle\rho_{hh,\,vv}$.

Applying a set of filters having $n_n = 2$ and using (28), the backscattering coefficients of the mean CM estimate resulting in maximum *a posteriori* probability are biased by $13/(13 + 2 \times 3) \approx -1.65$ dB. Average covariance matrices have been estimated for the eight areas marked in Fig. 3(a) for the original and the restored data, and Table I shows the bias estimates. The restored backscattering coefficients are all lower than the original backscattering coefficients, but the bias is smaller than that predicted by (28). Also, there is some variation of the bias factors
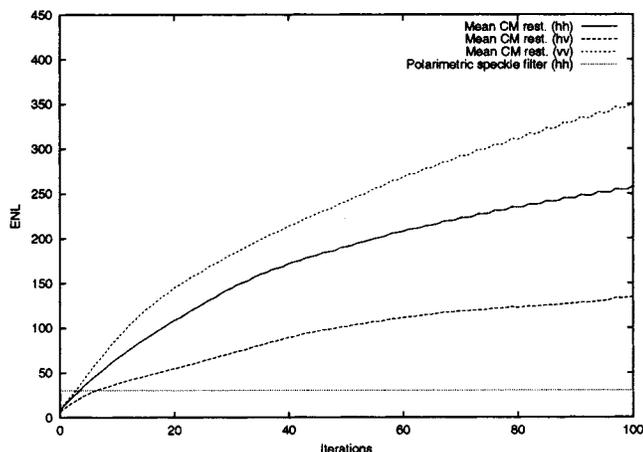
Fig. 5. Estimated equivalent number of looks for homogeneous field no. 1 for the mean CM restoration (all three polarizations) and for the noniterative polarimetric speckle filter (hh polarization).
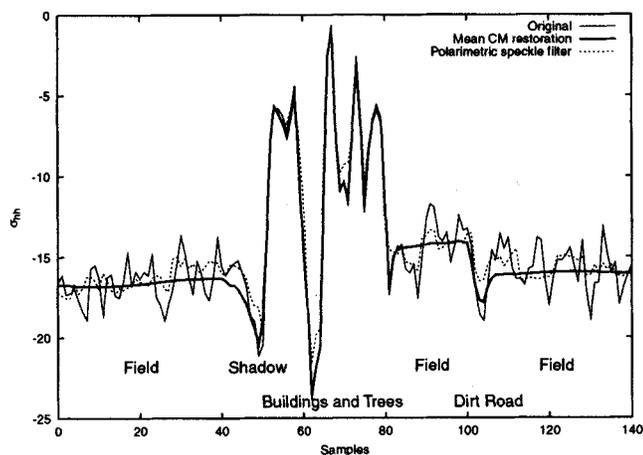

Fig. 6. Profile for hh polarization for the original data for the mean CM restoration and for the polarimetric SAR speckle filter.

TABLE I
BIAS OF RESTORED COVARIANCE MATRIX

Area	Crop	r_{hh}[dB]	r_{hv}[dB]	r_{vv}[dB]	$\Delta_{|\rho_{hh,vv}|}$	$\Delta_{\angle \rho_{hh,vv}}$ [deg]
1	Rye	-0.463	-0.295	-0.427	-0.022	-0.15
2	Oats	-0.493	-0.278	-0.434	-0.035	0.31
3	Winter wheat	-0.530	-0.439	-0.470	-0.020	-0.81
4	Spring barley	-0.462	-0.364	-0.496	-0.032	0.00
5	Grass	-0.382	-0.312	-0.439	-0.012	-1.40
6	Forest	-0.318	-0.509	-0.469	-0.007	1.68
7	Forest	-0.523	-0.443	-0.483	0.008	-0.46
8	Water	-0.530	-0.335	-0.566	-0.030	0.11

V. UNSUPERVISED CLASSIFICATION USING RESTORED DATA

The unsupervised classification scheme by Cloude and Pottier [15] examines the coherency matrix $\mathbf{T}$, which can be formed directly from the complex scattering amplitudes, or by a linear transformation of the covariance matrix [28]. According to the decomposition by Cloude [28], the coherency matrix is composed of three independent scattering mechanisms. The importance of each mechanism is found from an eigenvalue analysis of $\mathbf{T}$. Furthermore, from the eigenvalues λ_i, the parameter H is defined as

$$ H = \sum_{i=1}^{3} -P_i \log_3 P_i, \qquad P_i = \frac{\lambda_i}{\sum_{j=1}^{3} \lambda_j}. \qquad (29) $$

H is called the entropy and provides information about the randomness of the target. $H \in [0, 1]$ with $H = 0$ for a simple target (single scattering mechanism) and $H = 1$ for a random target. The α angle is derived from the associated eigenvectors, and α equals 0 for surface scattering $\pi/4$ for dipole scattering and $\pi/2$ for double-bounce scattering.

The entropy and the mean α angle $\overline{\alpha} = \sum_{i=1}^{3} P_i \alpha_i$ form the basis of the unsupervised classification scheme by Cloude and Pottier [15]. The H–α classification space is divided into nine zones based on general properties of the scattering mechanisms in the respective zones, making the classification approach data independent and truly unsupervised.

As discussed in Section III-D, the restored CM is biased, and as the coherency matrix is obtained by a linear transform of the CM, the eigenvalues of $\mathbf{T}$ will be biased as well. If the bias of the CM is given by a scaling of the true mean CM with a real scalar as suggested by (28), the eigenvalues are all biased by the same factor as the restored CM. In this case, the entropy is not influenced by this bias, as only the relative values of the eigenvalues are used for finding the entropy. The α angles are unaffected as well, as the eigenvectors are not affected by this bias, and hence, the classification is insensitive to a scaling of the CM with a real scalar. As shown in Section IV-E, however, the backscattering coefficients in the restored mean CM are biased with slightly different factors, and this affects the estimate of the entropy as well as the mean α angle.

To obtain a sufficiently high degree of multilooking, we apply the unsupervised classification to polarimetric SAR data filtered by the polarimetric speckle filter as done by Lee *et al.* [16] and to mean CM restored data, respectively. Fig. 7 shows the resulting entropy for the two cases, and we see that especially the CM restored data results in highly filtered estimates for the homogeneous areas. Fig. 8(a) shows the classification space using the

for the different polarizations within the respective fields. The complex correlation coefficient is practically unbiased for all the areas. The areas all have coherences above 0.45, except the forest areas having $|\rho_{hh,vv}| \approx 0.3$ and grass having $|\rho_{hh,vv}| = 0.07$. If the comparisons are based on the entire scene, we obtain $r_{hh} = -0.522$ dB, $r_{hv} = -0.500$ dB, $r_{vv} = -0.519$ dB, $\Delta_{|\rho_{hh,vv}|} = -0.010$, and $\Delta_{\angle \rho_{hh,vv}} = -0.11$.

This analysis has shown that the backscattering coefficients in the restored mean CM are not biased by exactly the same factor. The factors are in the same range, however, and the complex correlation coefficient remains practically unbiased, thus justifying the discussion in Section III-D. The bias of the restored CM in the homogeneous areas is significantly lower than the bias expression in (28), but this expression is for the ideal case, where the n_n neighbors within the connection filter all are given by $\mathbf{Z}_{jk_0} = [L(\kappa - p)/\kappa(L + n_n p)]\Sigma_0$. In practice, we sample the complex Wishart distribution using a very limited set of filters. This means that (28) cannot be used directly for correction of the bias, rather, it justifies that bias is indeed introduced by the algorithm. If the bias is compensated using the bias estimates from the entire scene, the backscattering coefficients are generally biased less than 0.2 dB in the homogeneous areas.
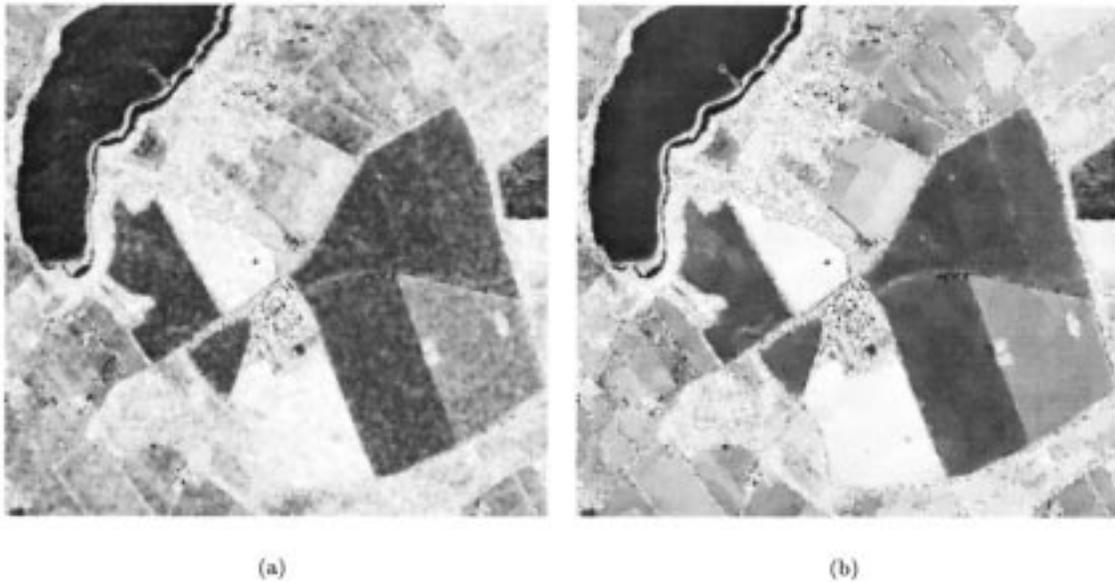
Fig. 7.   Entropy from the Cloude decomposition using filtered polarimetric data. (a) Using polarimetric SAR speckle filtered data and (b) using mean CM restored data.
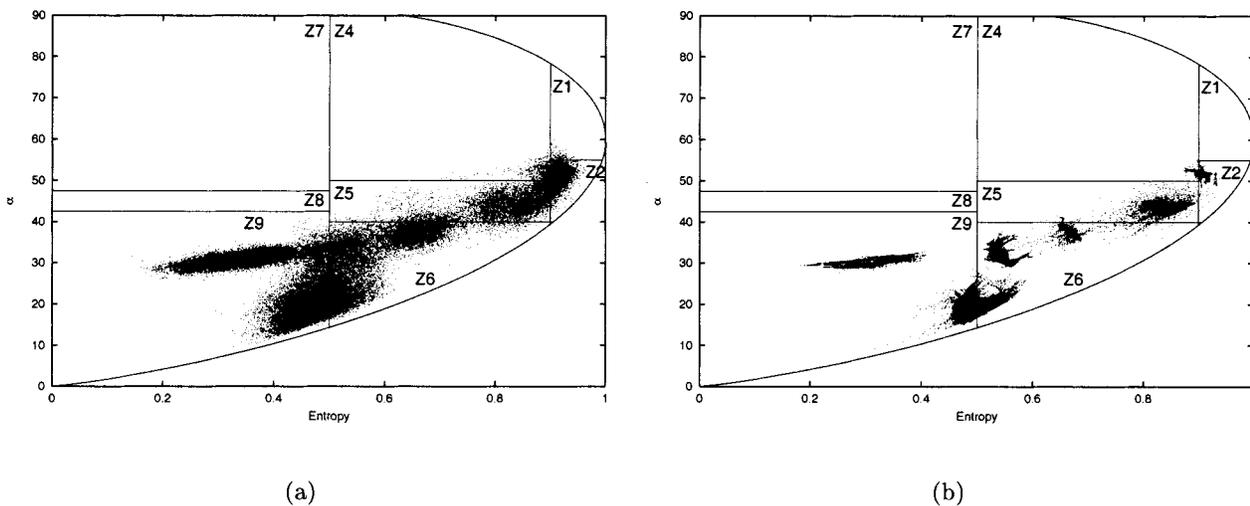

(a)   (b)

Fig. 8. $H-\alpha$ classification space including the zones Z1: High entropy multiple scatter; Z2: High entropy vegetation scattering; Z4: Medium entropy multiple scattering; Z5: Medium entropy vegetation scattering; Z6: Medium entropy surface scattering; Z7: Low entropy multiple scattering; Z8: Low entropy dipole scattering; and Z9: Low entropy surface scattering. Also, $H-\alpha$ distribution for eight land-use areas. (a) Using polarimetric SAR speckle filtered data and (b) using mean CM restored data.

same boundaries as Cloude and Pottier [15]. The $H-\alpha$ distribution for the eight land-use areas using the polarimetric speckle filtered data are shown as well in the figure, and we see that the distributions generally overlap. In Fig. 8(b), the distribution using the mean CM restored data is shown, and it is noticeable, that most of the land use classes are now clustered and easily separated due to the significantly lower variance of the individual distributions. To investigate the bias of the entropy and mean alpha angle, the center of each land use class has been estimated. Table II shows the result when applying original data and mean CM restored data, and we observe maximum deviations of 0.024 for H and 0.95° for $\overline{\alpha}$. The polarimetric speckle filter does not result in bias.

In Fig. 8 it is seen that the $H-\alpha$ distribution from water is in the low entropy surface scattering zone as expected at L-band

TABLE II
AVERAGE VALUES OF H AND α ESTIMATED FROM THE ORIGINAL AND THE MEAN CM RESTORED DATA

Area	Crop	H_{mean}	α_{mean}	H_{rest}	α_{rest}
1	Rye	0.486	19.30	0.505	20.24
2	Oats	0.523	32.45	0.547	33.40
3	Winter wheat	0.490	19.03	0.505	19.56
4	Spring barley	0.659	37.93	0.673	38.71
5	Grass	0.915	50.87	0.918	51.11
6	Forest	0.858	43.42	0.852	43.46
7	Forest	0.832	44.48	0.837	44.78
8	Water	0.331	30.69	0.350	31.42

[15]. Several of the crops also have surface scattering characteristics, as the radar signal penetrates through the crops to the surface below. This behavior is most pronounced for rye and winter wheat, and it has been confirmed by investigation of the
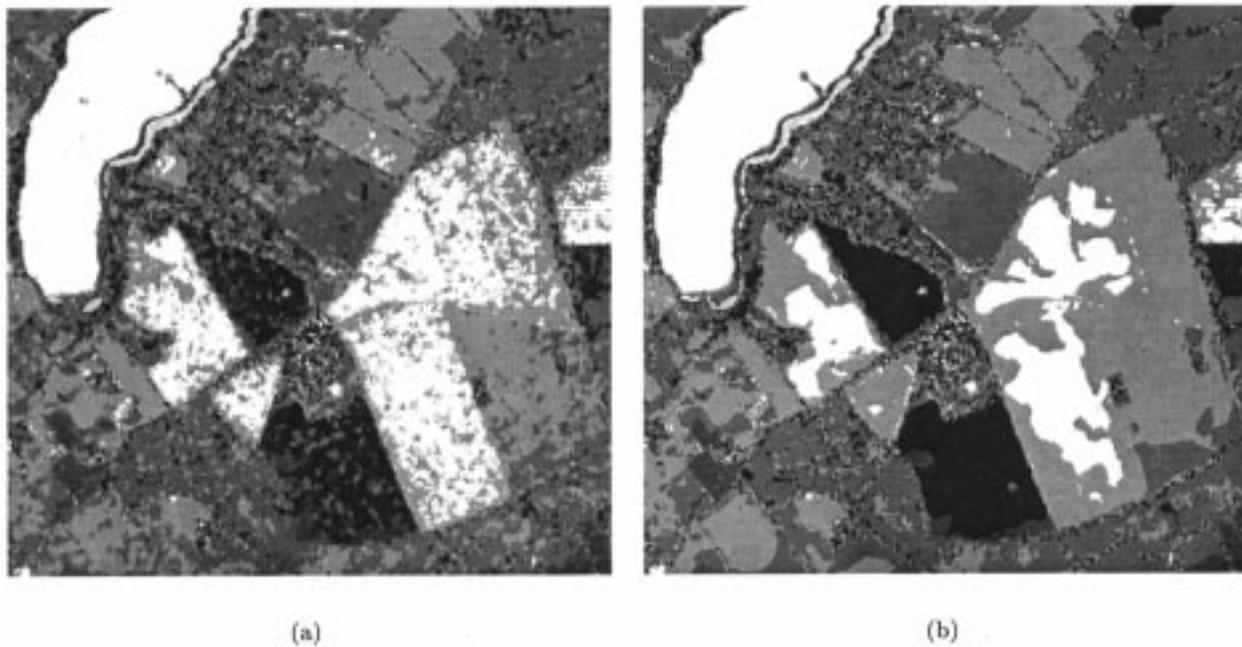
Fig. 9. Classification map using unsupervised classification. (a) Using polarimetric SAR speckle filtered data and (b) using mean CM restored data.

polarization responses [29]. Grass has a much larger entropy compared to the other crops due to the totally random orientation of the blades. Forest lies primarily in the medium entropy vegetation zone located at the high end of the entropy range, indicating that volume scattering from the canopy is a substantial contribution.

Some of the distributions in Fig. 8(a) and (b) are centered at or near the boundary between two zones, and this is reflected in the classification maps shown in Fig. 9. As expected from Figs. 7 and 8 the result using the mean CM restored data appears more homogeneous due to the clustered distributions. Still, the structures are preserved in the image. Also, it is quite interesting that, in the case of mean CM restored data, the classification map for the rye and winter wheat fields, both appearing as a mixture between two scattering mechanisms in Fig. 8, has connected areas within the field which is either low- or medium entropy surface scattering rather than having a more random appearance. This suggests, that it could be related to actual physical characteristics of, e.g., the soil. The unsupervised classification scheme in combination with the restoration algorithm proves in general to be an effective and consistent method for partitioning the area into classes having distinct scattering mechanisms, thus assisting the interpretation of the scene.

VI. CONCLUSION

In this work, a novel approach for mean CM estimation is presented, capable of handling multifrequency and/or multitemporal polarimetric SAR data. By incorporating the statistics of the polarimetric SAR data in the restoration and applying a set of small filters in an iterative scheme, the algorithm provides highly filtered data for the homogeneous areas while maintaining most of the structures in the image. The algorithm assumes homogeneous surfaces, but the approach should be applicable for textured surfaces as well by including a texture

variable in the complex Wishart distribution describing the polarimetric SAR data. In the optimization we apply simulated annealing, which is normally characterized by its slow convergence rate, but when using SAR data having a relatively high number of looks, only a rather limited number of iterations is necessary to obtain highly filtered results. By subdividing the set of filters the computational load of the algorithm is reduced by a factor of three, approximately, thus making the restoration algorithm feasible. The restoration introduces radiometric distortion, and an analytical expression of the bias is presented. The expression explains the bias in an ideal case and is not directly applicable for compensating for the bias, but using an average bias estimate, the homogeneous areas are biased less than 0.2 dB. The restored mean CM contains the full polarimetric information and can thus be used by any application using polarimetric data. It has been successfully applied as a preprocessing step for an unsupervised classification scheme, resulting in a quite homogeneous classification result. Also, it can be used in combination with segmentation and target-detection algorithms.

REFERENCES

[1] Ulaby and Elachi, *Radar Polarimetry for Geoscience Applications*. New York: Wiley, 1985.
[2] L. M. Novak and M. C. Burl, "Optimal speckle reduction in polarimetric SAR imagery," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 26, pp. 293–305, Feb. 1990.
[3] A. Lopès and F. Séry, "Optimal speckle reduction for the product model in multilook polarimetric SAR imagery and the Wishart distribution," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 632–647, May 1997.

[4] G. Liu, S. Huang, A. Torre, and F. Rubertone, "The multilook polarimetric whitening filter (MPWF) for intensity speckle reduction in polarimetric SAR images," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1016–1020, May 1998.

[5] J. S. Lee, M. R. Grunes, and S. A. Mango, "Speckle reduction in multipolarization, multifrequency SAR imagery," *IEEE Trans. Geosci. Remote Sensing*, vol. 29, pp. 535–544, July 1991.

[6] J. S. Lee, M. R. Grunes, and G. de Grandi, "Polarimetric SAR speckle filtering and its implication for classification," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2363–2372, Sept. 1999.

[7] C. Oliver and S. Quegan, *Understanding Synthetic Aperture Radar Images*. New York: Artech House, 1998.

[8] R. G. White, "A simulated annealing algorithm for radar cross-section estimation and segmentation," *Proc. SPIE*, vol. 2243, pp. 231–240, 1994.

[9] ——, "A simulated annealing algorithm for SAR and MTI image cross-section estimation," *Proc. SPIE*, vol. 2316, pp. 137–145, 1994.

[10] I. McConnell, R. White, C. Oliver, and R. Cook, "Radar cross-section estimation of SAR images," *Proc. SPIE*, vol. 2584, pp. 164–175, 1995.

[11] I. McConnell and C. Oliver, "Comparison of annealing and iterated filters for speckle reduction in SAR," *Proc. SPIE*, vol. 2958, pp. 74–85, 1996.

[12] D. Geman, *Random Fields and Inverse Problems in Imaging*. Berlin, Germany: Springer-Verlag, 1990, vol. 1427.

[13] S. Z. Li, *Markov Random Field Modeling in Computer Vision*. Berlin, Germany: Springer-Verlag, 1995.

[14] J. S. Lee and M. R. Grunes, "Classification of multi-look polarimetric SAR data based on complex Wishart distribution," in *Nat. Telesystems Conf.*, Washington, DC, 1992, pp. 7/21–24.

[15] S. R. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 68–78, Jan. 1997.

[16] J. S. Lee, M. R. Grunes, T. L. Ainsworth, L.-J. Du, D. L. Schuler, and S. R. Cloude, "Unsupervised classification using polarimetric decomposition and the complex Wishart classifier," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2249–2258, Sept. 1999.

[17] O. Hellwich, "Model Parameter Estimation using Simulated Annealing," in *Int. Arch. Photogrammetry and Remote Sensing, ISPRS Symp., Commission III*, vol. 32, Columbus, OH, 1998, pp. 233–238.

[18] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI–6, pp. 721–741, 1984.

[19] L. Ingber, "Adaptive simulated annealing (ASA): Lessons learned," *Contr. Cybern.*, vol. 25, no. 1, pp. 33–54, 1996.

[20] L. Tsang, J. A. Kong, and R. T. Shin, *Theory of Microwave Remote Sensing*. New York: Wiley, 1985.

[21] N. R. Goodman, "Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction)," *Ann. Math. Statist.*, vol. 34, pp. 152–177, 1963.

[22] S. H. Yueh, J. A. Kong, J. K. Jao, R. T. Shin, H. A. Zebker, and T. Le Toan, "K-distribution and multi-frequency polarimetric terrain radar clutter," *J. Electromagn. Waves Applicat.*, vol. 5, no. 1, pp. 1–15, 1991.

[23] M. Brookes. Matrix Reference Manual, Available: http://www.ee.ic.ac.uk/hp/staff/dmb/matrix/ intro.html. Imperial College, London, U.K.

[24] N. J. Higham and H.-M. Kim, "Numerical analysis of a quadratic matrix equation," Dept. Math., Manchester Centre Comput. Math., Univ. Manchester, Aug. 1999.

[25] E. L. Christensen, N. Skou, J. Dall, K. W. Woelders, J. H. Jrgensen, J. Granholm, and S. N. Madsen, "EMISAR: An absolutely calibrated polarimetric L- and C-band SAR," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1852–1865, Nov. 1998.

[26] H. Skriver, M. T. Svendsen, and A. G. Thomsen, "Multitemporal C- and L-band polarimetric signatures of crops," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 2413–2429, Sept. 1999.

[27] C. Addison, E. Appiani, R. Cook, M. Corvi, J. Harms, G. Howard, and B. Stephens, "PULSAR: Parallel noise despeckling of SAR images," *Lecture Notes Comp. Sci.*, vol. 1067, pp. 177–182, 1996.

[28] S. R. Cloude and E. Pottier, "A review of target decomposition theorems in radar polarimetry," in *IEEE Trans. Geosci. Remote Sensing*, Mar. 1996, vol. 34, pp. 498–518.

[29] J. J. van Zyl and F. T. Ulaby, "Scattering matrix representation for simple targets," in *Radar Polarimetry for Geoscience Applications*, F. T. Ulaby and C. Elachi, Eds. New York: Artech House, 1990.

Jesper Schou received the M.Sc. degree in engineering from the Technical University of Denmark (DTU), Lyngby, in 1997, where he is currently pursuing the Ph.D. degree in electrical engineering on a project dealing with feature extraction from SAR images for cartographic mapping.

His primary research interest is image analysis, with a special emphasis on SAR images.

Henning Skriver received the M.Sc. and Ph.D. degrees in 1983 and 1989, respectively, both in electrical engineering, from the Technical University of Denmark (DTU), Lyngby.

He has been with the Section of Electromagnetic Systems (EMI), DTU, since 1983, where he is now an Associate Professor. His work has primarily been concerned with various topics related to the utilization of SAR data for different applications. From 1983 to 1992, his main area of interest was retrieval of sea ice parameters from SAR data, including SAR data from ERS-1. Since 1992, he has covered different aspects of land applications of SAR data, such as forestry in the MAESTRO-1 project, agricultural and environmental applications using both satellite SAR data and data from the Danish airborne polarimetric SAR, EMISAR. His interests also include various methods for processing of SAR data, such as SAR image simulation, SAR image filtering, speckle statistics, texture analysis, segmentation, calibration, and polarimetric analysis. Most recently, he has been a Project Manager of a project concerning the use of SAR data for cartographic mapping.