

Article

A Hierarchical Spatio-Temporal Markov Model for Improved Flood Mapping Using Multi-Temporal X-Band SAR Data

Sandro Martinis * and André Twele

German Aerospace Center (DLR), German Remote Sensing Data Center (DFD), Oberpfaffenhofen, 82234 Wessling, Germany; E-Mail: andre.twele@dlr.de

* Author to whom correspondence should be addressed; E-Mail: sandro.martinis@dlr.de; Tel.: +49-8153-283-034; Fax: +49-8153-281-445.

Received: 28 July 2010; in revised form: 10 September 2010 / Accepted: 10 September 2010 /

Published: 17 September 2010

Abstract: In this contribution, a hybrid multi-contextual Markov model for unsupervised near real-time flood detection in multi-temporal X-band synthetic aperture radar (SAR) data is presented. It incorporates scale-dependent, as well as spatio-temporal contextual information, into the classification scheme, by combining hierarchical marginal posterior mode (HMPM) estimation on directed graphs with noncausal Markov image modeling related to planar Markov random fields (MRFs). In order to increase computational performance, marginal posterior-based entropies are used for restricting the iterative bi-directional exchange of spatio-temporal information between consecutive images of a time sequence to objects exhibiting a low probability, to be classified correctly according to the HMPM estimation. The Markov models, originally developed for inference on regular graph structures of quadtrees and planar lattices, are adapted to the variable nature of irregular graphs, which are related to information driven image segmentation. Entropy based confidence maps, combined with spatio-temporal relationships of potentially inundated bright scattering vegetation to open water areas, are used for the quantification of the uncertainty in the labeling of each image element in flood possibility masks. With respect to accuracy and computational effort, experiments performed on a bi-temporal TerraSAR-X ScanSAR data-set from the Caprivi region of Namibia during flooding in 2009 and 2010 confirm the effectiveness of integrating hierarchical as well as spatio-temporal context into the labeling process, and of adapting the models to irregular graph structures.

Keywords: Markov random fields (MRFs); hierarchical marginal posterior mode (HMPM) estimation; irregular graph; spatio-temporal context; TerraSAR-X; automatic thresholding; generalized Gaussian distribution; change detection; flood mapping; flood possibility mask; Namibia

1. Introduction

Spaceborne synthetic aperture radar (SAR) remote sensing is currently the most effective technology for a regular observation of the Earth's surface. By analyzing multi-date images, it is possible to get a synoptic view of temporal land-cover/land-use behavior. This is of particular importance in disaster management for the detection of abrupt changes caused by natural catastrophes like earthquakes, hurricanes and tsunamis, as well as for the monitoring of events characterized by a longer temporal evolution such as flooding and forest fires. SAR plays a particularly important role in flood mapping, due to its near all-weather/day-night capabilities and its effectiveness to detect inundation beneath different kinds of vegetation canopies [1-3].

A large portion of the activations (~46%, current as of August 19, 2010) of the International Charter of Space and Major Disasters (<http://www.disasterscharter.org>) are related to flood situations. At the same time, the number of SAR sensors that allow monitoring inundations with a high temporal and spatial resolution has strongly increased. For this reason, automatic methods are necessary to generate detailed flood maps on a near real-time basis to support decision makers as well as humanitarian relief organizations during disaster management. It can further be useful to create maps, which show the probability of correctness related to the classification results. This supplementary information source may be important for performance evaluations of hydrodynamic models.

Many techniques have been presented to identify change areas in remote sensing data [4,5]. The two most common image change detection algorithms are labeling of feature maps (e.g., difference images [6], normalized difference images [7], ratio data [8], and log ratio data [9]) and comparison of individual classification maps [10], usually called Post-Classification Comparison (PCC). In contrast to the classification of feature maps, PCC can provide a complete matrix of change directions. However, its performance naturally depends on the accuracy of each single classification map. It is therefore of great importance to integrate the temporal dimension into the classification process in order to improve results in terms of accuracy and reliability [11,12].

Markov random fields (MRFs) [13,14] represent an effective and theoretically well established probabilistic model for integrating different types of contextual information (e.g., spatial, hierarchical, and temporal) into the image labeling process. MRFs have also been adopted for an improved extraction of changes in remote sensing data. Spatial Markov modeling is applied to feature maps in several studies (e.g., [9,15,16]). A Markovian model considering spatial as well as hierarchical contextual information is presented by [7]. Several methods are proposed to integrate only temporal [12,17], as well as spatio-temporal [11,18,19], information into the Markov model for classification improvement.

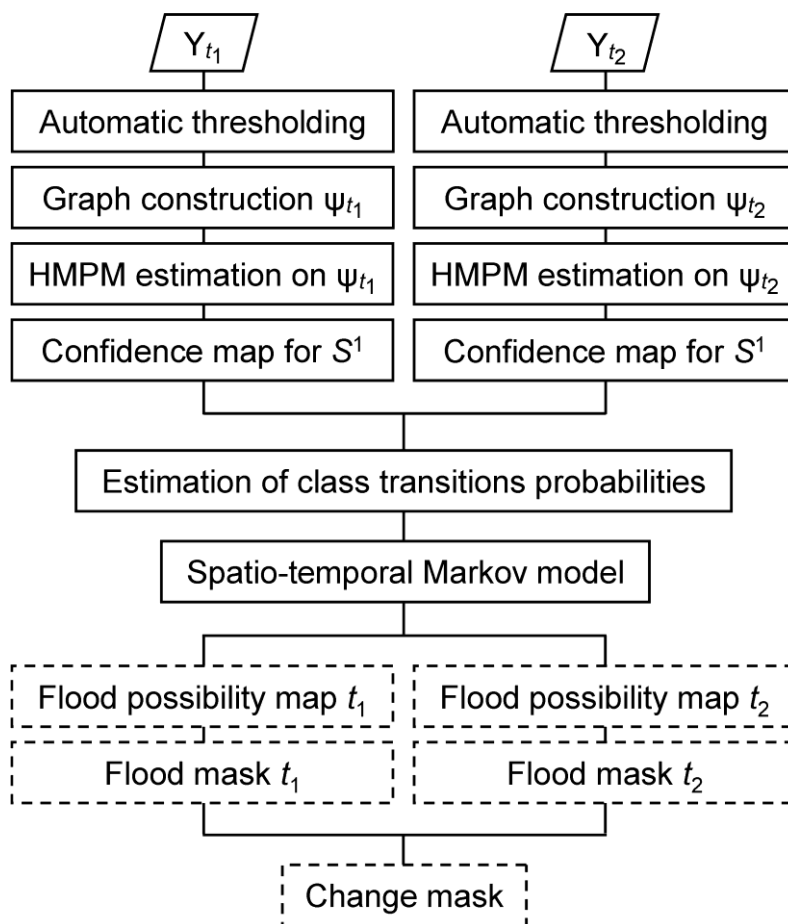
The application of MRFs in image analysis is predominantly accomplished by using noncausal (e.g., [9,11,16,19-24]) and planar causal models (e.g., [15,23,25]) on regular planar lattice-based graphs. These models have the disadvantage that they must be applied to all pixels in the image. This results in an enormous computational complexity, which is more pronounced using noncausal MRFs due to their iterative nature of inference, in contrast to the non-iterative character of planar causal models. In addition, only a small neighborhood is considered in the labeling process.

In contrast to noncausal or planar causal models, hierarchical causal Markov models have the capacity to capture the intrinsic hierarchical nature of remote sensing data. These models can be applied in a non-iterative way on simple regular structures of quadrees [26-31] or on more complex, however still regular, trees which try to overcome the blockiness of the classification result that is related to the nonstationarity of MRFs on quadrees [26,32].

In just a few works in image analysis, Markov modeling is related to irregular graphs, whose elements have a strong correlation with real objects or areas of the Earth's surface: Feitosa *et al.* [12] use planar graphs for multi-temporal classification based on fuzzy Markov chains. In [33], hierarchical marginal posterior mode (HMPM) estimation is applied to a Multiscale Region Adjacency Tree, however without consideration of the altering relationship between parent and child nodes in the graph architecture. This is taken into account by [7], who adapt hierarchical maximum *a posteriori* (HMAP) estimation to irregular graphs. This work also presents a framework for the supplementary integration of spatial context into the labeling process in a time-efficient manner. This is accomplished by using noncausal Markov modeling on a restricted region of the finest tree-level specified by confidence maps derived from the HMAP labeling result. Even if the integration of spatial context results in an increase in accuracy, the applied noncausal Markov model is not adapted to the irregular nature of the planar graph.

This paper presents an unsupervised method for improved near real-time flood and change detection in multi-temporal TerraSAR-X data using a hybrid generative Markov model (Figure 1). The specific objectives are: (1) the combination of automatically initialized causal and noncausal Markov image models for integrating hierarchical as well as spatio-temporal context into the labeling process of the classes "Water", "No water" and "Potentially flooded vegetation"; (2) the adaption of these models to irregular hierarchical as well as planar graphs; (3) the use of entropy-based confidence maps, derived in conjunction with hierarchical marginal posterior mode (HMPM) estimation, to restrict the noncausal Markov modeling step to regions that have been probably misclassified by the causal Markov model, and to generate flood possibility (FP) maps.

Figure 1. Block scheme of the proposed workflow.



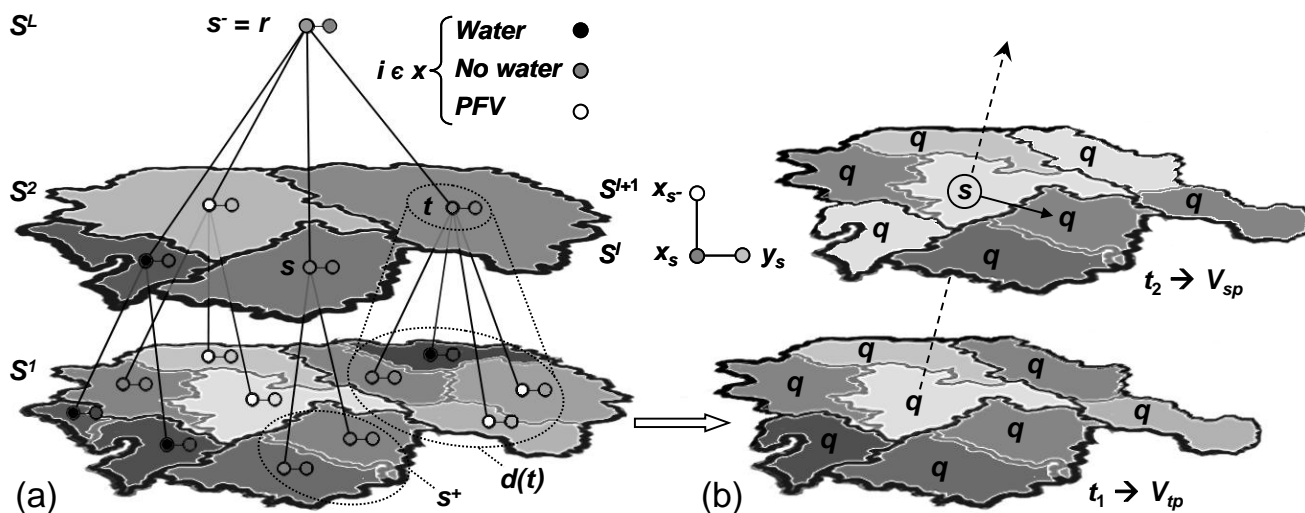
2. Methodology

2.1. Automatic Graph Construction

The multi-level representation of an image Y can be represented as a connected graph Ψ_L with L levels composed of a set of nodes S , where each node s apart from the root r has a unique parent node s' (Figure 2(a)). A descendant node originating from s is denoted as node t . The set of nodes can be partitioned into different scale levels, $S = S^1 \cup S^2 \cup \dots \cup S^L$. The level decomposed by the lowest homogeneity parameter is S^1 , and the coarsest level consists of only one node $S^l = \{r, l = L\}$ (Figure 2(a)).

Several differences can be determined between the regular structure of quadtrees and irregular hierarchical graphs: in contrast to the predefined architecture of quadtrees, which are represented by four equally sized child objects per parent object, the number and size of child objects per irregular parent object vary as a result of an information-driven segmentation. Therefore, the number of nodes of the whole graph cannot be estimated before image segmentation is accomplished. This is related to problems of over- or under-segmentation. Considering the intra-scale dependencies, it can be stated that, in contrast to regular planar lattices, the size and number of the elements vary, as well as the border length between the center and the adjacent objects in a spatial neighborhood system of an irregular planar graph. These characteristics must be taken into account when generating the irregular graph and when applying the Markov models to this tree structure.

Figure 2. (a) Three-level independency graph corresponding to an irregular tree-structure. Left circles represent labeling of classes “Water”, “No water” and “Potentially flooded vegetation” (PFV), right circles represent the observations at each node s . (b) Irregular planar graphs of Y_{t_1} and Y_{t_2} with spatial V_{sp} and temporal V_{tp} neighborhood systems of center node s and its spatially and temporally adjacent objects q at t_2 .



For the automatic generation of the hierarchical graph according to user demands, the procedure described in [7] is adopted: First, several subsets of the SAR image are automatically selected to describe the heterogeneity of the SAR data. In the next step, a pre-segmentation of the subsets is performed by the fractal net evolution approach [34]. The homogeneity parameter is estimated, which leads to a decomposition of the entire image with average object sizes of the segments at each level, which come close to those intended by the user. This is accomplished by generating a database, which contains models describing the relationship between homogeneity parameter and object size according to data of different SAR sensor types and image contents. Finally, this model is selected for creating the whole graph which best fits to the pre-segmentation result.

2.2. Markov Image Modeling

Context is an essential information source for the analysis of remotely sensed data. Its necessity gains in importance with increasing spatial resolution of the data. The use of context means that each image element is not treated in isolation but as part of a spatial pattern. MRFs have been found to be useful tools for integrating different types of context into the classification process. In the following, two different Markov models are described: A causal model using a hierarchical marginal posterior mode (HMPM) estimation on two separated irregular hierarchical graphs, and a noncausal model leading to a bi-directional exchange of spatio-temporal information between multi-temporal planar graphs, represented by the finest levels of the hierarchical tree. Additionally, a hybrid model is proposed, which combines the HMPM estimation with the noncausal model for integrating hierarchical as well as spatio-temporal context into the labeling process in a time efficient manner using confidence maps.

2.2.1. Causal Markov Modeling on Irregular Graphs

Problem Definition and Statistical Modeling

For the set of nodes of an irregular hierarchical graph, two sets of random variables $X = (X_s)_{s \in S}$, $Y = (Y_s)_{s \in S}$, called “random fields” are considered. Each X_s takes its values in a finite set of M classes $\Omega = \{\omega_1, \omega_i, \dots, \omega_M\}$ and each Y_s takes its values in \mathfrak{R} . As common in Bayesian modeling, the labeling problem is then to estimate the “best” unobserved realization $X = x$ given the observed realization $Y = y$, where $y = (y_s)_{s \in S}$ is the observed image element.

Assuming a first-order Markov chain on a directed independence graph, where the conditioning for each node in $S^l (l \neq R)$ reduces to a dependence on its parent node in S^{l+1} , as well as a standard site-wise factorization for the observation model $P(y|x)$, the joint distribution factorizes as a product of local functions [27]:

$$P(x, y) = P(x_r) \prod_{s \neq r} P(x_s | x_{s^-}) \prod_{s \in S} P(y_s | x_s) \quad (1)$$

where $P(x_r)$ are the root prior probability, $\{P(x_s | x_{s^-})\}_{s \neq r}$ are the parent-child transition probabilities and $\{P(y_s | x_s)\}_{s \in S}$ are the data conditional likelihoods.

Model Parameters

For the definition of the prior model $P(X_s | X_{s^-})$ on the graph, the Potts-like distribution is used in [26], which favors likeliness of the labeling at s and s^- . In this work, the computation of an inter-scale transition probability matrix is proposed. This matrix contains the transition probability of each class combination between levels S^l and S^{l+1} ($l \neq R-1$) as the ratio of the area assigned to class ω_j at S^l and ω_i at S^{l+1} among the area of all elements assigned to class ω_j at S^l . At root r , a uniform prior probability with a value of $1/M$ is chosen.

The conditional likelihoods $P(y_s | x_s)$ of each node are modeled as Gaussian mixtures, defined by the scale dependent parameter vector $\theta_i^l = (\mu_i^l, \sigma_i^l)$, with mean μ_i^l and standard deviation σ_i^l of class i at level l . For the estimation of the initial class statistics, an automatic parametric tile-based thresholding procedure [7,35] under the generalized Gaussian assumption is used to estimate threshold values τ_1 between the classes “Water” and “No water” as well as τ_2 between the classes “No water” and “Potentially flooded vegetation (PFV)” in large-size SAR data with small class *a priori* probabilities in a time efficient manner. The threshold computation is accomplished by using a generalized Gaussian Kittler and Illingworth (GG-KI) thresholding algorithm [36,37] on a small number of tiles of the entire image Y selected according to the probability of the tiles to contain a bi-modal mixture distribution of the classes to be separated. The computed values of τ_1 and τ_2 are then applied to each level of the graph to derive θ_i^l .

Inference

One of the most interesting aspects of this model lies in the possibility to reach exact inference of the labels by computing the maximum *a posteriori* (MAP) $P(x|y)$ and marginal posterior mode (MPM)

probabilities $P(x_s|y)$ using an extension of the Viterbi algorithm [38] as well as the Baum algorithm [39] on Markov chains in scales [27,28]. The cost function associated to the MAP criterion is

$$C(\hat{x}, x') = 1 - \delta(\hat{x}, x') \tag{2}$$

with estimated and true class label \hat{x} and x' , and the Kronecker delta function δ . Equation (2) penalizes the discrepancies between configurations without considering how different these configurations are. The cost function

$$C(\hat{x}, x') = \sum_{s \in S} \delta(\hat{x}_s, x'_s) \tag{3}$$

accounts for this aspect and leads to the following Bayesian MPM estimator, which associates the most probable class given all the data to each node (Table 1):

Table 1. HMPM estimation on irregular hierarchical graphs.

Preliminary pass: At this downward recursion, the marginal priors $P(x_s)$ are computed for each s : $s \in S^{L-1} \dots S^2$, $P(x_s) = \sum_{x_{s^-}} P(x_s | x_{s^-}) P(x_{s^-})$

Bottom-up sweep: The distribution of each x_s and couple (x_s, x_{s^-}) given all the data of the descendants (incl. y_s) is provided:
 Initialization ($s \in S^1$):
 $P(x_s | y_s) \propto P(y_s | x_s) P(x_s)$
 $P(x_s, x_{s^-} | y_s) = \frac{P(x_s | x_{s^-}) P(x_{s^-}) P(x_s | y_s)}{P(x_s)}$

Recursion ($s \in S^2 \dots S^L$):
 $P(x_s | y_{d(s)}) \propto P(y_s | x_s) P(x_s) \prod_{t \in s^+} \sum_{x_t} \left(\frac{P(x_t | y_{d(t)}) P(x_t | x_s)}{P(x_t)} \right) \phi_{HMPM}$
 $P(x_s, x_{s^-} | y_{d(s)}) = \frac{P(x_s | x_{s^-}) P(x_{s^-}) P(x_s | y_{d(s)})}{P(x_s)} \quad \forall s \neq r$

Top-down sweep: The complete marginal posteriors are reassembled from the partial marginals computed at the bottom-up sweep:
 Initialization (r):
 $P(x_r | y) = P(x_r | y_{d(r)})$
 $\hat{x}_r = \arg \max_{x_r} P(x_r | y)$

Recursion ($s \in S^{L-1} \dots S^1$):
 $P(x_s | y) = \sum_{x_{s^-}} \frac{P(x_s, x_{s^-} | y_{d(s)}) P(x_{s^-} | y)}{\sum_{x_s} P(x_s, x_{s^-} | y_{d(s)})}$
 $\hat{x}_s = \arg \max_{x_s} P(x_s | y)$

The weighting factor ϕ_{HMPM} takes into account the proportion of the spatial extent of a child node s to its predecessor s^- during the computation of the site-wise partial marginal posteriors $P(x_s | y_{d(s)})$ within the bottom-up pass.

2.2.2 Noncausal Markov Modeling on Bi-Temporal Planar Graphs

Spatio-Temporal Markov Model

Noncausal Markov models are, in contrast to hierarchical approaches, able to incorporate intra-spatial context between adjacent nodes of an undirected graph as well as temporal information into the labeling process. In the following, the pixel-based spatio-temporal Markov model proposed in [11] is reformulated and adapted to irregular planar graphs. In contrast to a uni-directed “cascade” approach [40] analyzing a sequence of images in chronological order, this model accomplishes an iterative mutual exchange of information between a sequence of multi-temporal images.

For the sake of simplicity, a bi-temporal sequence is considered consisting of two registered images Y_{t_1} and Y_{t_2} , which are acquired at times t_1 and t_2 , respectively. Each image is represented by an irregular hierarchical graph and labeled according to the HMPM estimation. For the application of the spatio-temporal model, the lowest levels of both graphs are synchronized by intersecting the two segmentations in order to retrieve equal object borders at t_1 and t_2 (Figure 2(b)).

The sets of possible labels for the related planar graphs are $\mathcal{A} = \{\lambda_1, \lambda_i, \dots, \lambda_N\}$ at t_1 and $\mathcal{Q} = \{\omega_1, \omega_i, \dots, \omega_M\}$ at t_2 . In the following, the labeling of Y_{t_2} , given Y_{t_1} and \mathcal{A} is considered. The optimal labeling of all image elements according to the MAP decision criterion is characterized by an enormous computational complexity. The processing demand can be substantially reduced by modeling the conditional prior probability of image elements within a local neighborhood associated with a clique system, which is a subset of adjacent image elements in a neighborhood system. According to the Hammersly-Clifford theorem [41], describing MRFs-Gibbs equivalence, the combination of the MAP estimation with MRFs makes the classification task equivalent to the minimization of the Gibbs energy function U expressed in the following relation:

$$P(x_s | y_s, X_{sp}, X_{tp}) = Z^{-1} \exp^{-U(x_s, y_s, X_{sp}, X_{tp})} \quad (4)$$

where Z is a normalizing factor called the partition function and X_{sp} and X_{tp} are subsets of the labels of the respective images in predefined spatial V_{sp} and temporal V_{tp} neighborhood systems. The Iterated Conditional Modes (ICM) algorithm [42] represents a computationally moderate solution of the MRF-MAP estimates, converging to a local, but usually sufficient minimum of Equation (4). Assuming that the contribution of spatial and temporal context is separable and additive, the Gibbs energy function to be minimized for node s by the ICM algorithm can be formulated as

$$U(x_s, y_s, X_{sp}, X_{tp}) = U_{data}(\cdot) + \gamma_{sp} U_{sp}(\cdot) + \gamma_{tp} U_{tp}(\cdot) \quad (5)$$

where γ_{sp} and γ_{tp} control the influence of spatial and temporal context into the flood detection process, respectively. Assuming conditional independence of y given x , the data term U_{data} can be formulated as:

$$U_{data}(X_s = \omega_i, Y_s = y_s) = -\ln(P(Y_s = y_s | X_s = \omega_i)) \quad (6)$$

The single-time posterior probabilities are modeled as Gaussian mixtures using class statistics derived from the HMPM estimation on the irregular graph. The spatial U_{sp} and temporal U_{tp} energy functions can be expressed as:

$$U_{sp}(X_s = \omega_i, X_{sp}) = - \sum_{\omega_j \in \Omega_{sp}} \varphi_{sp} \delta(\omega_i, \omega_j) \left(\sum_{\omega_j \in X_{sp}} \delta(\omega_i, \omega_j) \frac{\text{card}(q \in V_{sp-sim})}{\text{card}(q \in V_{sp})} \right) \quad (7)$$

and

$$U_{tp}(X_s = \omega_i, X_{tp}) = - \sum_{\lambda_j \in \Lambda_{tp}} \varphi_{tp} P(\omega_i | \lambda_j) \left(\sum_{\lambda_j \in X_{tp}} \delta(\omega_i, \lambda_j) \frac{\text{card}(q \in V_{tp-sim})}{\text{card}(q \in V_{tp})} \right) \quad (8)$$

Index δ is the Kronecker delta function, which is, according to a pair-wise clique system (s - q) between center node s and a spatially or temporally adjacent object q (Figure 2(b)), equal to 1 if $\omega_i = \omega_j$ or $\omega_i = \lambda_i$, respectively, and zero otherwise. Due to the fact that, in contrast to lattice-based graphs, the number of neighbors in V_{sp} and V_{tp} varies on irregular planar graphs, the amount of the spatial and temporal energy may vary significantly in relation to the spectral term. To equalize the influence of the spectral and temporal terms in the ICM estimation over the whole planar graph, an n^{th} -order spatial neighborhood system V_{sp-sim} is simulated, so that for example a first-order system contains four elements q . The temporal neighborhood system V_{tp-sim} is identical to V_{sp-sim} . However, since it considers the central object, it contains one additional element (Figure 2(b)).

In order to account for the varying size of the objects in the irregular neighborhood system V_{tp} , the weighting factor φ_{tp} is integrated, considering the spatial proportion of the respective classes in V_{tp} :

$$\varphi_{tp} = (A(\lambda_i, X_{tp}) / A(V_{tp})) \quad (9)$$

In addition, the weighting of the spatial energy function by φ_{sp} accounts for the varying border length between object s and q in the irregular neighborhood system V_{sp} :

$$\varphi_{sp} = ((A(\omega_i, X_{sp}) / A(V_{sp})) + b_{\omega_i}) / 2 \quad \text{with } b_{\omega_i} = \sum_{q \in \omega_i} b(s, q) / b_s \quad (10)$$

where b_{ω_i} describes the ratio of the shared border length of objects labeled to class ω_i with an adjacent object q to the total border length b_s of s . The greater the size of an object q and the higher the common border to s , the more weight is given to this object at the computation of U_{sp} .

The same formulations of the MRF related to Equations (6–10) can be generated for image Yt_1 . The mutual approach couples the two MRFs of the bi-temporal data-set using the ICM algorithm. Starting from an initial labeling produced by the HMPM step, the preceding classification of one image is exploited to update the labeling of the other image at each iteration, and *vice versa*, until convergence. For an extension of this bi-directional approach to a sequence of more than two images, the reader is referred to [11].

Estimation of Temporal Transition Probabilities

The only term that considers correlation between data in (8) is $P(\omega_i | \lambda_i)$. It represents the transition probability from class λ_i at t_1 to class ω_i at t_2 . It is an element of the so-called transition probability matrix (TPM), which can be calibrated from expert's knowledge using ground truth or historic data [11] or estimated directly from the image data using for example the "compound classification rule" [43,44]. This method involves finding the optimal pair of classes (ω_i, λ_i) for each pair of image elements based on the Bayes rule for minimum error. In this work, the method of [44] is used, which replaces the TPM by a joint probability matrix (JPM), and is adapted to an irregular graph structure. The method uses an

iterative fixed-point EM (Expectation-Maximization) -like algorithm [45] for estimating the temporal correlation of multi-date images represented by joint class probabilities $P(\omega_i, \lambda_i)$. The elements of the related JPM of size $N \times M$ are computed recursively for k iterations until convergence is reached according to:

$$P_{k+1}(\omega_i, \lambda_i) = \frac{1}{P(\omega_i)P(\lambda_i)O} \sum_{s=1}^O \frac{P_k(\omega_i, \lambda_i)P(\omega_i|y_s)P(\lambda_i|y_q)\varphi_{jpm}}{\sum_{\omega_n \in \Omega} \sum_{\lambda_m \in \Lambda} \frac{P_k(\omega_n, \lambda_m)}{P_k(\omega_n)P_k(\lambda_m)} P(\omega_n|y_s)P(\lambda_m|y_q)} \quad (11)$$

$P_0(\omega_i, \lambda_i)$ is initialized by $(N \times M)^{-1}$ and O is the number of objects of the planar graph. The *a priori* probabilities $P(\omega_i)$ and $P(\lambda_i)$ are derived by computing the relative occurrence of each class according to the respective HMPM estimation. The factor $\varphi_{jpm} = A_s/A_{\text{card}(s)}$ considers the irregular nature of the planar graph by weighting the influence of each object according to its relative spatial extent in relation to the entire scene. The following normalization generates values of the JPM, summing up to 1:

$$P_{k+1}^{\text{norm}}(\omega_i, \lambda_i) = \frac{P_{k+1}(\omega_i, \lambda_i)}{\sum_{\omega_n \in \Omega} \sum_{\lambda_m \in \Lambda} P_{k+1}(\omega_n, \lambda_m)} \quad (12)$$

2.2.3. Hybrid Multi-Contextual Markov Model

The knowledge of marginal posteriors $P(x_s|y)$ allows to access the degree of confidence for each node s by computing the marginal posterior entropy E_s [28]:

$$E_s(x_s|y) = - \sum_{\omega_i \in \Omega} P(x_s = \omega_i|y) \log P(x_s = \omega_i|y) \quad (13)$$

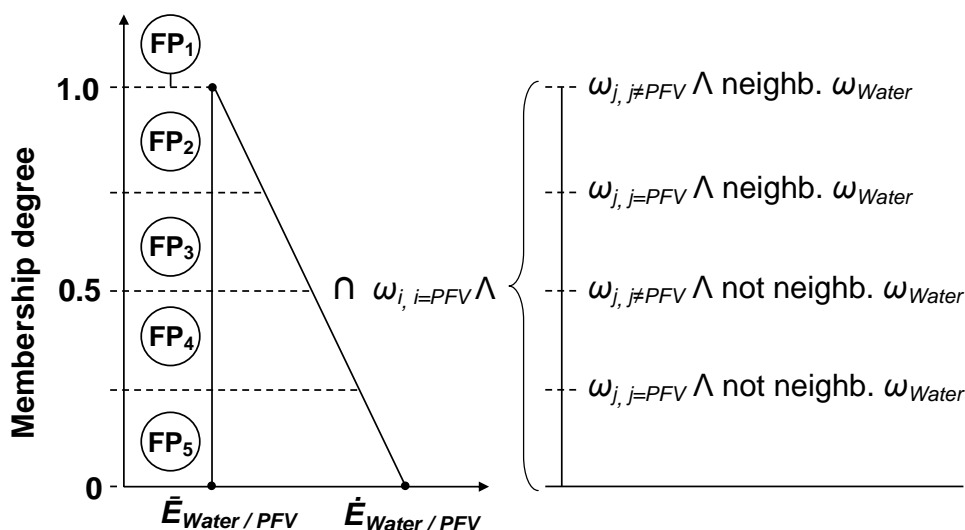
Based on this formulation, a confidence map can be computed for the whole graph. Higher entropy values are indicators of misclassifications of s [46]. These confidence maps are used for the combination of the hierarchical causal and the spatio-temporal noncausal Markov model by restricting the ICM algorithm presented in (4) to values of E_s greater than a defined threshold T_E . This decreases the computational effort since the spatio-temporal context is only integrated for the label estimation of a limited number of elements of the sub-area $S_{\text{sub-1}}^1$. After the first iteration, this number can be further reduced by restricting the ICM algorithm to objects $s \in S_{\text{sub-1}}^1$ with a common border to other elements greater T_E .

2.3. Generation of Flood Probability Maps

A quantification of uncertainty in the labeling of each image element can be valuable for flood disaster management and performance evaluation of hydrodynamic models. Recently, Schumann *et al.* [47] used inundation possibilities derived from multi-algorithm ensembles for calibrating flood models. In this work, entropy-based confidence maps are used created from the final labeling result of ω_{Water} and ω_{PFV} for the generation of a five-class flood possibility (FP) map. Fuzzy theory is used to transform the entropy values to fuzzy memberships in the interval $[0, \dots, 1]$ (Figure 3). Full weight (FP₁) is given to nodes $s \in \omega_{\text{Water}}$ and $s \in \omega_{\text{PFV}}$ with E_s lower than the mean entropy \bar{E}_{ω_i} of ω_{Water} and ω_{PFV} , respectively. Objects with entropies between \bar{E}_{ω_i} and the maximum entropy \hat{E}_{ω_i} are assigned to four different

possibility classes (FP_{2,5}) according to a linearly decreasing fuzzy membership function. The fuzzy set of ω_{PFV} is combined with additional membership values taking into account spatial as well as temporal relationships of $s \in \omega_{PFV}$ from the final classification; e.g. full weight is given to objects with $E_s \leq \bar{E}_{PFV}$, labeled to a different class at Δt and neighboring elements $s \in \omega_{Water}$. In contrast, elements without contact with other flood objects and exhibiting no class change over time are given a lower degree of membership, since these objects exhibit a reduced likelihood of being inundated (Figure 3).

Figure 3. Fuzzy sets for the generation of FP maps using MPM entropies and spatio-temporal labeling information.



3. Experimental Results

3.1. Data-set Description

In order to carry out an experimental analysis aimed at assessing the performance of the proposed approach for a real test case, a bi-temporal data-set consisting of two TerraSAR-X ScanSAR scenes with 8.25 m pixel spacing in range and azimuth direction is used (Figure 4(a,b)). The study area of this work is the Zambezi floodplain situated in the Caprivi Strip in north-eastern Namibia, which is regularly affected by flooding related to heavy seasonal rainfalls. Both images, which were acquired on April 11, 2009 and April 20, 2010, show large scale inundations with a higher flood level in 2009. Since both data-sets were acquired in an identical orbit with similar ordering options (HH-polarization, incidence angle range of 27.1 °–36.6 °), no major system-related differences are present.

In both data-sets, in comparison to the surrounding dry land, open water areas appear dark due to specular reflection of the incident radar signal. In contrast, flooded vegetation causes very distinct and bright signatures. Indeed, X-band SAR has a strongly reduced ability to detect inundation beneath dense vegetation such as forest due to increased canopy attenuation and volume scattering in comparison to the longer C-, and L-band signals (e.g., [48]). In this study area, however, the emergent vegetation is mainly composed of foliated shrubs and grassland, whose structure admits a multiple-bounce effect, in which the penetrated radar pulse is backscattered from the water surface and lower sections of the vegetation. This causes a high signal return [49]. For simplicity, most flood

mapping algorithms only consider open water areas. This, however, results in an underestimation of the flood extent if the flood plain is interspersed with emergent vegetation. Digital elevation models (DEMs) have been used effectively for detecting flooding beneath vegetation [35,50]. However, since the study area is characterized by a low topographic variability, and a high resolution DEM was not at our disposal within this study, no such information could be used to optimize classification results.

For validation purposes of the water extent in each SAR scene, two optical images from the RapidEye satellite of April 14, 2009 and April 22, 2010, with a spatial resolution of 5 m, were available. The time-offset relative to the SAR data was three days and two days, respectively. However, due to the stable flood conditions, no critical change in the flood extents was observable between the SAR and optical data-sets. For this reason, the RapidEye scenes were used to create a reference map for each date for a small sub-area (256×256 pixel) of the eastern part of Lake Liambezi by visual interpretation and manual digitalization of the open water and flooded vegetation areas. High resolution optical images of historic flood events were available in Google Earth for cross-checking the validation mask.

3.2. Results and Discussion

In this section, the effectiveness of the proposed multi-contextual hybrid Markov model is evaluated by investigating the influence of (1) using irregular graphs, (2) adapting the Markov models to these irregular structures, and (3) combining hierarchical causal with noncausal Markov modeling on irregular graphs. Additionally, the results of the HMPM and HMAP estimation are compared.

For this purpose, several classification results were generated for the chosen reference areas of Y_{t_1} and Y_{t_2} by applying different models to regular and irregular graphs (Table 2). R-T-ICM considers the pixel-based integration of only spatio-temporal information into the labeling process. The sole integration of hierarchical context is accomplished by R-HMAP and R-HMPM on a quadtree as well as by I-HMAP and I-HMPM on an irregular graph, without considering the weighting of the child objects in relation to their parent nodes. The non-weighted hybrid Markov model is represented by R-HMAP-ICM and R-HMPM-ICM on a quadtree as well as I-HMAP-ICM and I-HMPM-ICM on an irregular graph. In contrast, I-HMAP-ICM-w and I-HMPM-ICM-w consider the variable structure of hierarchical and spatio-temporal neighborhood systems of irregular graphs.

Identical threshold values of $\tau_1 = 50.7$ and $\tau_2 = 50.0$ for Y_{t_1} as well as $\tau_1 = 134.1$ and $\tau_2 = 163.2$ for Y_{t_2} derived by the automatic thresholding approach mentioned in Section 2.2.1 are used for the initialization of each model. The regularization parameters γ_{sp} and γ_{tp} are fixed to 1 during the experiment. A real (regular graph) and simulated (irregular graph) first-order neighborhood system is used in the noncausal Markov model, which is only applied to objects with an entropy value greater than the average entropy \bar{E} of Y_{t_1} and Y_{t_2} , respectively. The hierarchical model at times t_1 and t_2 is accomplished on graphs with eight levels. Therefore, on a quadtree, the smallest objects are defined on S^1 by 2×2 pixels. Accordingly, S^1 is partitioned into 16,384 nodes. Due to the fixed decrease of 25% of the object number between adjacent scales, the total number of nodes of the tree is 21,845. The irregular graphs are built with a relative object number of ~50% between S^{l+1} and S^l . The finest levels of Y_{t_1} and Y_{t_2} are partitioned into ~2,000 homogeneous segments. This results in a total object number

of the respective graphs of ~4,300. After synchronizing the segmentations of the two planar graphs, both levels S^1 are represented by ~12,000 nodes.

In order to assess the performance of the different classifiers, the ‘‘Overall Accuracy’’ (OA) measure is used, which is the percentage of correctly classified pixels in relation to the total number of pixels (Table 2).

The experimental results show a nearly identical mean OA of 79.5% for Y_{t_1} and 80.1% for Y_{t_2} , averaged over all algorithms (Table 2). The non-hierarchical R-T-ICM method provides OA similar to these average values and significantly outperforms the HMAP, HMPM and HMAP-ICM estimation on the quadtree as well as the HMAP-w estimation on the irregular graph. However, the R-T-ICM method is accompanied by a strong computational demand, since it is iteratively applied to every pixel in the images within six iterations until convergence is reached.

Table 2. Overall accuracy of different models according to graph structure chosen. Index ‘‘w’’ marks methods which consider a weighting in the classification according to the irregular nature of the graphs.

| Time | Regular (R-) | | | | | Irregular (I-) | | | | | | Mean |
|-------|-----------------|----------|-------|--------------|--------------|----------------|------------|----------------|----------------|--------------|--------------|------|
| | Planar T-ICM | Quadtree | | | | Hierarchical | | | | | | |
| | | HMAP | HMPM | HMAP- ICM | HMPM- ICM | HMAP- w | HMPM- w | HMAP-ICM- w | HMPM-ICM- w | HMAP- ICM | HMPM- ICM | |
| t_1 | 79.90 | 77.08 | 78.50 | 77.62 | 80.30 | 77.80 | 79.67 | 80.39 | 82.97 | 78.63 | 81.78 | 79.5 |
| t_2 | 79.77 | 73.80 | 76.97 | 78.11 | 81.04 | 76.23 | 80.86 | 82.48 | 86.33 | 80.07 | 85.20 | 80.1 |

It can clearly be seen that the results established with the HMPM estimates show higher OAs than methods that are related to the HMAP approach. The difference is ~1.4% to 2.7% for Y_{t_1} and more distinct for Y_{t_2} with values between ~2.6% and 5.1%. Among the two hierarchical estimators, the HMPM estimator is more time consuming due to the higher computational complexity in the top-down pass. However, it offers the possibility to compute confidence maps based on MPM entropies.

In comparison to the hierarchical modeling results, the OA is consistently higher when taking into account the bi-directional exchange of spatio-temporal information within the ICM algorithm. The increase in accuracy is in the range of ~1.0% to ~2.5% for Y_{t_1} and significantly higher for Y_{t_2} with values of 3.3% to 6.8%.

If one considers the graph structure in the multi-contextual hybrid Markov modeling result, it is notable that the non-weighted models I-HMAP-ICM and I-HMPM-ICM are more efficient for Y_{t_2} with OAs ~2.0% and ~4.2% higher than the quadtree-based models R-HMAP-ICM and R-HMPM-ICM. In contrast, for Y_{t_1} , the respective OAs are only increased by ~1.0% and ~1.5%. By adapting the hybrid models (I-HMAP-ICM-w and I-HMPM-ICM-w) to the irregular graph structures, a further increase in accuracy can be stated for both images compared to the non-weighted models I-HMAP-ICM and I-HMPM-ICM. Highest OAs of ~83.0% for Y_{t_1} and ~86.3% for Y_{t_2} are both related to the proposed method HMPM-ICM-w, which offers an accuracy increase of ~2.7% and ~5.3%, respectively, compared to the quadtree-based model R-HMPM-ICM. This can be explained by reduced discrepancies in the labeling process using homogeneous irregular objects instead of quadratic segments, which are characterized by a higher probability to contain mixtures of different classes.

Additionally, in contrast to noncausal Markov modeling on a regular planar lattice, spatially larger neighborhood systems can be considered. This is associated with a fast convergence of the ICM algorithm within four iterations. Besides the increased accuracy, the computational demand of the hierarchical models on irregular graphs is reduced by a factor of about five in comparison to quadtrees due to the lower number of nodes (~4,000 vs. ~21,845). Likewise, the application of the ICM algorithm is characterized by less computational complexity on irregular planar graphs due to the reduced number of nodes at S^1 (~12,000 nodes) in comparison to the finest levels (16,384 nodes) of the quadtrees, despite the fact that additional weighting parameters must be computed.

When applied to the sub-area of Y_{t_1} and Y_{t_2} , the proposed multi-contextual Markov model takes 35 s of processing time on two irregular hierarchical graphs with eight levels and a total number of respectively ~4,300 nodes (CPU details: Intel Xeon 5460 Core Duo @ 3.16-GHz processor with a 3-GB RAM). The generation of the hierarchical graph and the HMPM estimation account for ~45%, the estimation of the temporal transition probabilities and the application of the ICM estimator require ~55% of this time. The execution of the proposed method on irregular hierarchical graphs is ~60% faster than on quadtrees with eight levels, even if the generation of the irregular graph is more time consuming and a weighting of the nodes must be taken into account.

The incorporation of contextual information has a smoothing effect on the final classification results, by screening out isolated segments of a certain class in homogeneous regions (Figure 4(c,d)). Errors mainly occur due to an underrepresentation of the open water area at the land-water boundaries. This is due to the coarser resolution of the ScanSAR data in contrast to the RapidEye reference data and the high sensibility of the radar signal to protruding cyperaceous vegetation at the river-banks, which prohibits the perceptibility of a distinct waterline. As is shown in Table 3, this effect is more pronounced at Y_{t_1} (UA: 78.9%) than at Y_{t_2} (UA: 93.2%), where smaller water areas are present. However, producer accuracy (PA) reaches more than 93% for both dates. Highest errors are related to the misclassification of image elements of class “PFV” to “No water”. These errors are mainly caused by flooded vegetation areas with a dense canopy, which prohibits multiple-bounce effects. This results in an attenuated signal return and an underestimation of the class “PFV”. This effect is more pronounced at t_2 with an UA of class “PFV” of ~50.0% which is ~20.0% lower compared to that at t_1 .

If one combines the classes “Water” and “PFV” to one single class “Flood”, the supplementary consideration of bright scatterers into the labeling process enhances flood mapping accuracy (UA) from 33.9% to 73.8% at t_1 and from 62.2% to 78.9% at t_2 . The large difference of the UA of class “Flood” between both dates results from the fact that the proportion of open water areas is significantly higher at t_2 due to a higher flood level. This causes a complete submersion of protruding vegetation areas which act as strong scatterers at t_1 . The main flood areas exhibit a high probability to be classified correctly according to the FP maps (Figure 4(e,f)). Smaller, more isolated open water areas are marked by a higher classification uncertainty given the HMPM estimation.

Presumably, the application of the proposed method to C- or L-band SAR data would result in a higher detectability of flooding beneath vegetation. However, as the contrast between open water and land areas is most distinct in X-band SAR data, a lower classification accuracy of the class “Water” may be expected using SAR systems of greater wavelength.

Figure 4. TerraSAR-X data-sets for Caprivi/Namibia (center coordinate: 17°56'44"S, 24°22'34"E) on (a) April 11, 2009 and (b) April 20, 2010; Classification results derived by using the IR-HMPM-ICM-w model (c, d); FP maps (e, f).

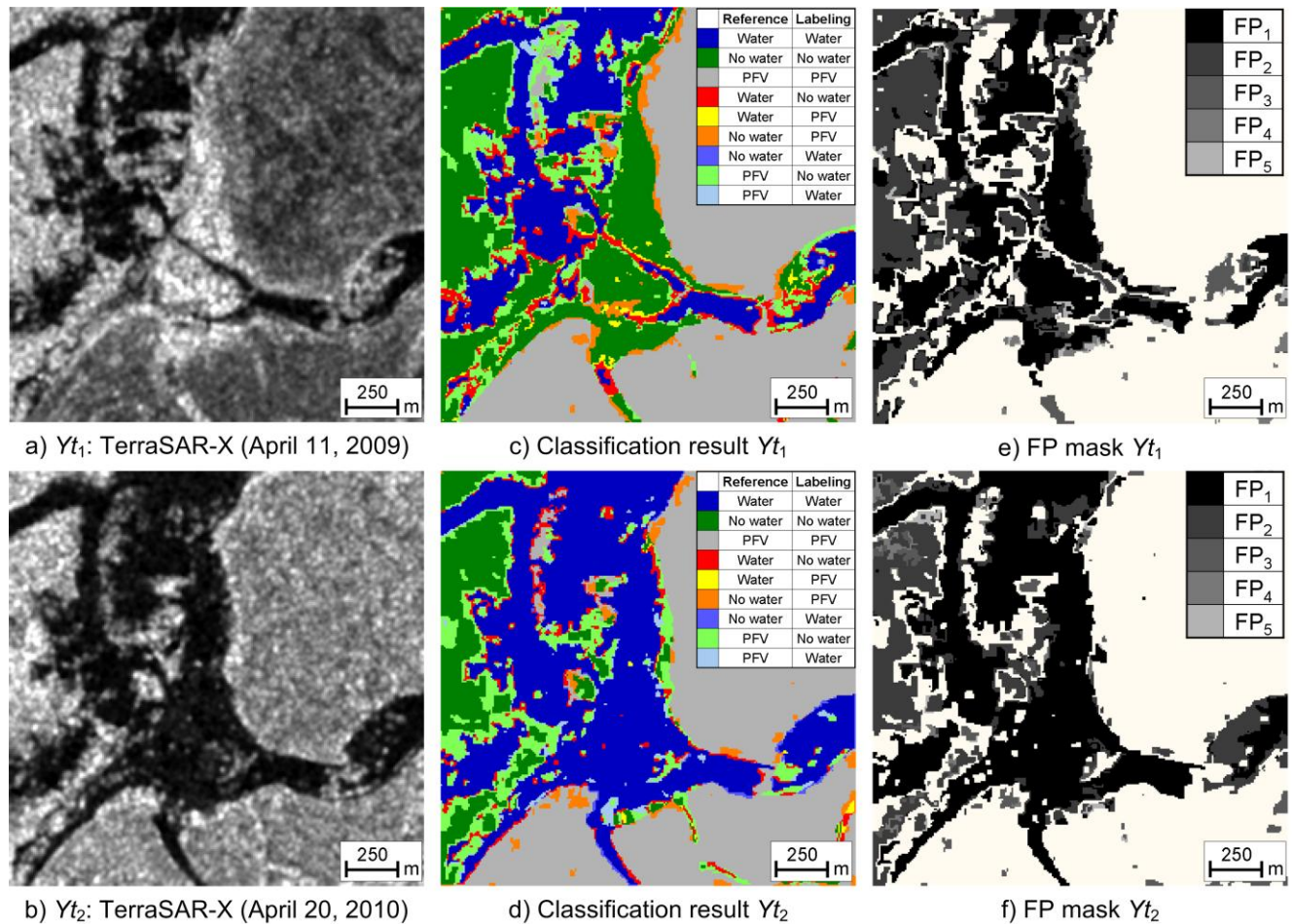


Table 3. Accuracy matrices for the labeling result of HMPM-ICM-w on an irregular graph at t_1 and t_2 (PA = Producer Accuracy, UA = User Accuracy).

| Date | Classified | Reference | | | UA [%] | |
|------|-----------------|-----------|----------|-------|-----------------|-------|
| | | Water | No water | PFV | | |
| T1 | Water | 11129 | 2857 | 115 | 78.92 | |
| | No water | 541 | 30261 | 1895 | 92.54 | |
| | PFV | 236 | 5511 | 12991 | 69.32 | |
| | | 11906 | 38629 | 15001 | 65536 | |
| | PA [%] | 93.47 | 78.33 | 86.60 | OA [%] | 82.97 |
| | | | | | Khat [%] | 71.69 |
| T2 | Water | 21240 | 1486 | 61 | 93.21 | |
| | No water | 145 | 29673 | 1546 | 94.61 | |
| | PFV | 744 | 4976 | 5665 | 49.67 | |
| | | 21610 | 38038 | 5888 | 65536 | |
| | PA [%] | 95.98 | 82.18 | 77.90 | OA [%] | 86.33 |
| | | | | | Khat [%] | 77.20 |

4. Conclusion

In this work, the problem of unsupervised flood detection in SAR data using a hybrid multi-contextual Markov image model on irregular hierarchical graphs is addressed. Experimental results obtained on a bi-temporal TerraSAR-X ScanSAR data-set for Caprivi/Namibia confirm the effectiveness of combining automatically initialized causal hierarchical with noncausal spatio-temporal Markov models with respect to accuracy and computational performance.

The incorporation of hierarchical context into the labeling process is accomplished by a hierarchical marginal posterior mode (HMPM) estimation using Markov chains in scale. The supplementary integration of spatio-temporal context is applied by an iterative bi-directional information exchange between elements of two consecutive planar graphs of a time sequence. This processing step is restricted to elements which offer a certain probability to be wrongly classified according to marginal posterior-based entropies.

It is shown that the application of this hybrid model on irregular graphs helps to reduce classification errors and computational demands in comparison to modeling on the regular graph structure of quadrees. The adaption of this model to the variable nature of irregular hierarchical and spatio-temporal neighborhood systems leads to a further increase in accuracy. Furthermore, it is demonstrated that HMPM estimation outperforms results generated by hierarchical maximum *a posteriori* (HMAP) estimation.

The incorporation of inundated bright scattering vegetation areas into the classification scheme results in a significant enhancement of flood mapping results. Entropy-based confidence maps combined with spatio-temporal relationships of potentially flooded vegetation to open water areas are used for the quantification of the uncertainty in the labeling of each image element in flood possibility masks.

For future work, additional information such as digital elevation models (DEMs) could be used to extend the fuzzy sets describing flood probabilities in areas with terrain of sufficient variability.

Acknowledgements

The major part of this work has been funded by the German Ministry of Education, Science, Research and Technology (BMBF) in the context of the Research Program “Risk Management of Extreme Flood Events” (RIMAX), project SAR-HQ (Grant Reference Number: 0330719). The authors would like to thank the two anonymous reviewers for their helpful comments.

References

1. Ormsby, J.P.; Blanchard, B.J.; Blanchard, A.J. Detection of lowland flooding using active microwave systems. *Photogramm. Eng. Remote Sens.* **1985**, *51*, 317-328.
2. Alsdorf, D.E.; Melack, J.M.; Dunne, T.; Mertes, L.A.K.; Hess, L.L.; Smith, L.C. Interferometric radar measurements of water level changes on the Amazon flood plain. *Nature* **2000**, *404*, 174-177.
3. Horrit, M.S.; Mason, D.C.; Cobby, D.M.; Davenport, I.J.; Bates, P. Waterline mapping in flooded vegetation from airborne SAR imagery. *Remote Sens. Environ.* **2003**, *85*, 271-281.

4. Lu, D.; Mausel, P.; Brindizio, E.; Moran, E. Change detection techniques. *Int. J. Remote Sens.* **2004**, *25*, 2365-2407.
5. Mas, J.F. Monitoring land-cover changes: A comparison of change detection techniques. *Int. J. Remote Sens.* **1999**, *20*, 139-152.
6. Bruzzone, L.; Prieto, D.F. Automatic analysis of the difference image for unsupervised change detection. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 1171-1182.
7. Martinis, S.; Twele, A.; Voigt, S. Unsupervised extraction of flood-induced backscatter changes in SAR data using Markov image modeling on irregular graphs. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, Accepted.
8. Moser, G.; Serpico, S.B. Generalized minimum-error thresholding for unsupervised change detection from SAR amplitude imagery. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 2972-2982.
9. Bruzzone, L.; Prieto, D.F. An adaptive semiparametric and context-based approach to unsupervised change detection in multitemporal remote-sensing images. *IEEE Trans. Image Process.* **2002**, *11*, 452-466.
10. Singh, A. Digital change detection techniques using remotely-sensed data. *Int. J. Remote Sens.* **1989**, *10*, 989-1003.
11. Melgani, F.; Serpico, S.B. A Markov random field approach to spatio-temporal contextual image classification. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 2478-2487.
12. Feitosa, R.Q.; Costa, G.; Mota, G.L.A.; Pakzad, K.; Costa, C.O. Cascade multitemporal classification based on fuzzy Markov chains. *ISPRS J. Photogramm. Remote Sens.* **2009**, *64*, 159-170.
13. Geman, S.; Geman, D. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **1984**, *6*, 721-741.
14. Li, S.Z. *Markov Random Field Modeling in Image Analysis*, 3rd ed.; Springer-Verlag: London, UK, 2009.
15. Carincotte, C.; Derrode, S.; Bourenane, S. Unsupervised change detection on SAR images using fuzzy hidden Markov chains. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 432-441.
16. Benedek, C.; Szirányi, T. Change detection in optical aerial images by a multilayer conditional mixed Markov model. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 3416-3430.
17. Mota, G.L.A.; Feitosa, R.Q.; Coutinho, H.L.C.; Liedtke, C.E.; Müller, S.; Pakzad, K.; Meirelles, M.S.P. Multitemporal fuzzy classification model based on class transition possibilities. *ISPRS J. Photogramm. Remote Sens.* **2007**, *62*, 186-200.
18. Solberg, H.S.; Taxt, T.; Jain, A.K. A Markov random field model for classification of multisource satellite imagery. *IEEE Trans. Geosci. Remote Sens.* **1996**, *34*, 100-113.
19. Liu, D.; Song, K.; Townshend, J.; Gong, P. Using local transition probability models in Markov random fields for forest change detection. *Remote Sens. Environ.* **2008**, *112*, 2222-2231.
20. Smits, P.; Dellepiane, S. Synthetic aperture radar image segmentation by a detail preserving Markov random field approach. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 844-857.
21. Jackson, Q.; Landgrebe, D. Adaptive Bayesian contextual classification based on Markov random fields. *IEEE Trans. Geosci. Remote Sens.* **2002**, *40*, 2454-2463.

22. Kasetkasem, T.; Varshney, K.V. An image change detection algorithm based on Markov random field models. *Canad. J. Remote Sens.* **2002**, *40*, 1815-1823.
23. Fjørtoft, R.; Delignon, Y.; Pieczynski, W.; Sigelle, M.; Tupin, F. Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 675-686.
24. Kasetkasem, T.; Arora, M.K.; Varshney, P.K. Super-resolution land cover mapping using a Markov random field based approach. *Remote Sens. Environ.* **2005**, *96*, 302-314.
25. Tso, B.; Olsen, R.C. Combining spectral and spatial information into hidden Markov models for unsupervised image classification. *Int. J. Remote Sens.* **2005**, *26*, 2113-2133.
26. Bouman, C.; Shapiro, M. A multiscale image model for Bayesian image segmentation. *IEEE Trans. Image Process.* **1994**, *3*, 162-177.
27. Laferté J.M.; Pérez, P.; Heitz, F. Discrete Markov image modeling and inference on the quadtree. *IEEE Trans. Image Process.* **2000**, *9*, 390-404.
28. Pérez, P.; Chardin, A.; Laferté J.M. Noniterative manipulation of discrete energy-based models for image analysis. *Pattern Recog.* **2000**, *33*, 573-586.
29. Cheng, H.; Bouman, C. Multiscale Bayesian segmentation using a trainable context model. *IEEE Trans. Image Process.* **2001**, *10*, 511-525.
30. Collet, C.; Murtagh, F. Multiband segmentation based on a hierarchical Markov model, *Pattern Recognit.* **2004**, *37*, 2337-2347.
31. Provost, J.-N.; Collet, C.; Rostaing, P.; Pérez, P.; Bouthemy, P. Hierarchical Markovian segmentation of multispectral images for the reconstruction of water depth maps. *Comput. Vis. Image Underst.* **2004**, *93*, 155-174.
32. Kato, Z.; Zerubia, J.; Berthod, M. A hierarchical Markov random field model and multi-temperature annealing for parallel image classification. *Graph. Models Image Process* **1996**, *58*, 18-37.
33. Katartzis, A.; Vanhamel, I.; Sahli, H. A hierarchical Markovian model for multiscale region-based classification of vector-valued images. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 548-558.
34. *eCognition Developer 8 User Guide*; Definiens AG: Munich, Germany, 2009.
35. Martinis, S.; Twele, A.; Voigt, S. Towards operational near real-time flood detection using a split-based automatic thresholding procedure on high resolution TerraSAR-X data. *Nat. Hazards. Earth Syst. Sci.* **2009**, *9*, 303-314.
36. Kittler, J.; Illingworth, J. Minimum error thresholding. *Pattern Recog.* **1986**, *19*, 41-47.
37. Bazi, Y.; Bruzzone, L.; Melgani, F. An unsupervised approach based on the generalized Gaussian model to automatic change detection in multitemporal SAR images. *IEEE Trans. Geosci. Remote Sens.* **2005**, *43*, 874-887.
38. Forney, G.D. The Viterbi algorithm. *IEEE Proc.* **1973**, *61*, 268-278.
39. Baum, L.; Petrie, T.; Soules, G.; Weiss, N. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *IEEE Ann. Math. Stat.* **2005**, *41*, 164-171.
40. Swain, P.H. Bayesian classification in a time-varying environment. *IEEE Trans. Syst. Man Cybern.* **1978**, *8*, 879-833.

41. Besag, J. Spatial interaction and the statistical analysis of lattice systems. *J. R. Stat. Soc. Ser. B* **1974**, *36*, 192-236.
42. Besag, J. On the statistical analysis of dirty pictures. *J. R. Stat. Soc. Ser. B* **1986**, *48*, 259-302.
43. Bruzzone, L.; Serpico, S.B. An iterative technique for the detection of land-cover transitions in multispectral remote sensing images. *IEEE Trans. Geosci. Remote Sens.* **1997**, *35*, 858-867.
44. Bruzzone, L.; Prieto, D.F.; Serpico, S.B. A neural-statistical approach to multitemporal and multisource remote-sensing image classification. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 1350-1359.
45. Granville, V.; Rason, J.P. Multivariate discriminant analysis and maximum penalized likelihood density estimation. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 501-517.
46. Feng, X.; Williams, C.K.I.; Felderhof, S.N. Combining belief networks and neural networks for scene segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 467-483.
47. Schumann, G.; Baldassarre, G.D.; Paul, P.D. The utility of spaceborne radar to render flood inundation maps based on multialgorithm ensembles. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 2801-2807.
48. Richards, J.A.; Woodgate, P.W.; Skidmore, A.K. An explanation of enhanced radar backscattering from flooded forests. *Int. J. Remote Sens.* **1987**, *8*, 1093-1100.
49. Hong, S.-H.; Wdowinski, S.; Kim, S.-W. Evaluation of TerraSAR-X observations for wetland InSAR application. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 864-873.
50. Mason, D.C.; Horritt, M.S.; Dall'Amico, J.T.; Scott, T.R.; Bates, P.D. Improving river flood extent delineation from synthetic aperture radar using airborne laser altimetry. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 3932-3943.

© 2010 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).