# Unsupervised Classification of Radar Imagery of Wetlands Using the Soft Competition Scheme

S. L. Durden and Z. S. Haddad

Jet Propulsion Laboratory

# 1. Introduction

- SAR has been shown to be sensitive to vegetation type and to presence or absence of surface water.
- Such information is useful in understanding the ecology of wetland areas and the impact of seasonal flooding.
- Vegetation type and presence of standing water are highly correlated with methane emission in boreal wetlands.
- Supervised classification has been used previously to distinguish land cover classes.
- Development of unsupervised method is highly desirable since training data is often not available.
- Method used here is a direct application of a technique developed for vector quantization.

## 2. Vector Quantization (VQ)

- VQ deals with the problem of finding a set of code vectors that can best represent data.
- This amounts to finding clusters in the input data and is identical to the unsupervised classification problem.
- Simplest clustering algorithm is k-means: assume class parameters and classify data, then re-compute class parameters.
- At each step, each input vector has one and only one class associated with it; i.e. "hard" classification.
- Better results may be obtainable using soft classification at each step: allow each input to be associated with multiple classes.

# 3. Soft Competition Scheme (SCS)

- SCS was developed by Yair, Zeger, and Gersho (1992) for performing vector quantization.
- We are presented with data samples $x(n)$, where $x$ is a vector representing the radar measurements at each pixel.
- The class mean vectors are $w_j$, where $j$ is the class index and ranges over the expected number of classes.
- We repeatedly examine the input data, updating $w_j$ according to

$$w_j(n+1) = w_j(n) + \eta_j(n) P_n(j)(x(n) - w(n))$$

- Here, $P_n(j)$ is the probability that $x(n)$ belongs to class $j$.
- This probability is estimated from $|x(n) - w_j(n)|$, the distance between the input vector and $w_j$.

$$P_n(j) = \frac{\exp(-\beta(n)|x(n) - w_j(n)|^2)}{\Sigma_k \exp(-\beta(n)|x(n) - w_k(n)|^2)}$$

- This is a Gibbs distribution where $\beta(n)$ is a parameter analogous to the inverse temperature.

# 4. Algorithm Details

- As algorithm is applied, $\beta$ is gradually raised, lowering temperature.
- Learning rate parameter $\eta_j$ varies with both the class and the time step.
- All $w_j$ are updated for each $x$, making the method "soft".
- $x$ that are closer to a particular $w$ have a larger effect on it during the update through the probability $P_n(j)$.
- Early in the procedure when $\beta$ is small, the probabilities are weak functions of the class $j$.
- This keeps the method from getting stuck in a local minimum.
- As $\beta$ increases, the probabilities become stronger functions of the class $j$.
- In the limit of very large $\beta$, the classifications become "hard".

## 5. Classification

- At completion of SCS, we have clusters of radar data.
- We can compute the statistics of each cluster; this is equivalent to the information used by a supervised classifier.
- For radar image classification we use the results of SCM as input to a Maximum *a Posteriori* (MAP) classifier.
- MAP maximizes $p(L|X)$ of the pixel labels $L$ conditioned on the radar observations $X$.
- We use Gaussian statistics for the PDF of the radar data conditioned on the class.
- A Markov Random Field model is used for the class PDF.
- Maximization is carried out using interated conditional modes (ICM).
- Final result of the method is a set of classes with known mean and variance and an image classified according to these class characteristics.

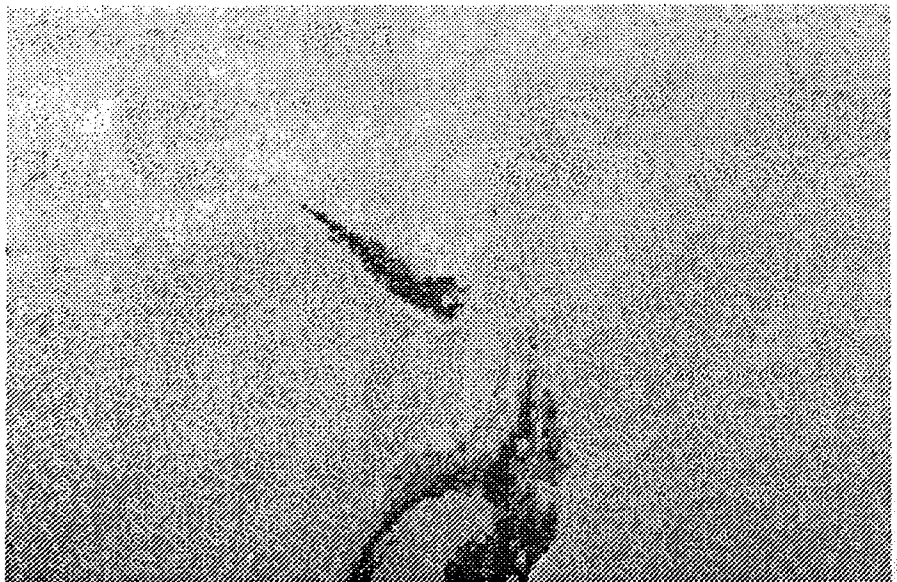
## 6. Results

- Method was applied to AIRSAR data collected over wetlands in Alaska and over Belize, Central America.
- For the AK data, the classifier uses LHH, LHV, CHH, and CHV.
- The AK data had been previously classified using a supervised approach (Durden et al. 1996) in which land cover classes correspond with methane production.
- For AK unsupervised and supervised agree well.
- For the Belize data the classifier input is HH at P-, L-, and C-bands.
- Several class numbers were tried; 4 appears to give best results.
- In Belize we do not consider methane production; we are simply interested in vegetation and surface water conditions.
- Blue areas likely correspond to flooded reeds (very low P-band return and very high C-band return); other areas are likely forest with varying degrees of surface water.

- Acknowledgment: The authors thank Drs. L. Morrissey and G. Livingston for their help in this work.

# Belize Wetlands Results

SAR PHH image



SCS Results