Parameter estimation and classification of multi-scale remote sensing data

G. Storvik
University of Oslo
Email: geirs@math.uio.no

R. Fjørtoft
Norwegian Computing Center (NR)
Email: Roger.Fjortoft@nr.no

A. H. S. Solberg
University of Oslo
Email: anne@ifi.uio.no

***Abstract*— In this article we describe a Bayesian model for integration of multi-scale image data. The approach is based on the concept of a reference resolution. Data at this and lower resolutions are connected to the reference resolution through a fully specified statistical model. Algorithms for parameter estimation and classification based on the multi-scale model are proposed, and results and comparisons with single-scale classification are presented.**

I. Introduction

The rapidly growing number of earth observation satellites provide a much better coverage in space, time and electromagnetic spectrum than in the past. Analysis of compound data sets therefore steadily gains importance. One of the key challenges in multi-sensor image fusion is how to combine images with different resolution to obtain more precise results.

The multi-scale model presented in this document is based on the concept of a reference resolution and is developed in a Bayesian framework [1]. We let the reference resolution correspond to the highest resolution present in the data set. For each pixel of the input image at the reference resolution we assume that there is an underlying discrete class. The observed pixel values are modeled conditionally on the classes, described at the reference resolution. The properties of the class label image are described by an *a priori* model. Markov random fields (MRFs) and Potts model have been selected for this purpose. Data at coarser resolutions are modeled as mixed pixels, i.e., the observations are allowed to include contributions from several distinct classes. In this way it is e.g. possible to exploit spectrally richer images at lower resolution to obtain more accurate classification results at the reference level, without losing as much details as if we simply oversampled the low resolution data prior to the analysis.

Multivariate Gaussian distributions are used to describe the data given the class. The classification is based on the Bayesian approach through maximization of the *a posteriori* probability, using the iterative conditional modes (ICM) algorithm.

The results of experiments on simulated images derived from SPOT XS and Landsat TM data are presented and discussed in view of the increased computing time.

A more detailed description with references is given in [2].

II. Model

The model consists of two parts, the prior model for the class image and the likelihood for data, which is a model conditional on the class image.

A. Model for class image

Define a reference resolution for which:

- Each pixel contain only one class (out of K possible).
- All observations are at this or a coarser resolution.

Assume the class image is described by $\mathbf{z} = (z_1, ..., z_n)$ where z_i defines the class in pixel i at the reference resolution. Consider Potts model

$$\pi(\mathbf{z}) = \frac{1}{C(\boldsymbol{\theta})} \exp\{\sum_i \alpha_{z_i} + \beta \sum_{i \sim i'} I(z_i = z_{i'})\} \qquad (1)$$

where $\boldsymbol{\theta} = (\alpha_1, ..., \alpha_K, \beta)$ and $C(\boldsymbol{\theta})$ is a normalization constant. $I(\cdot)$ is here the indicator function and $i \sim i'$ means that pixels i and i' are neighbors. This is an ordinary MRF model that is widely used for this purpose [3], [4].

B. Model for data

Assume that the observations $\mathbf{y} = (\mathbf{y}^1, ..., \mathbf{y}^p)$ are available, where $\mathbf{y}^j$ is a (possibly multi-channel) image at resolution r_j.

We furthermore assume that data at different resolutions are conditionally independent, i.e.

$$p(\mathbf{y}|\mathbf{z}) = \prod_{j=1}^{p} p(\mathbf{y}^j|\mathbf{z}).$$

This is a reasonable assumption for sensors with different spectral properties.

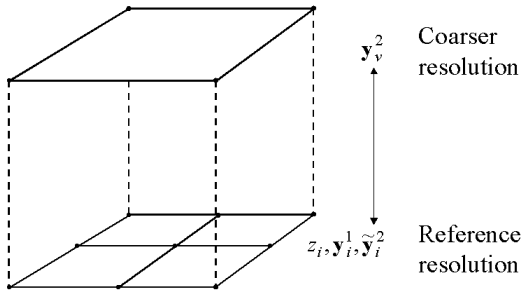
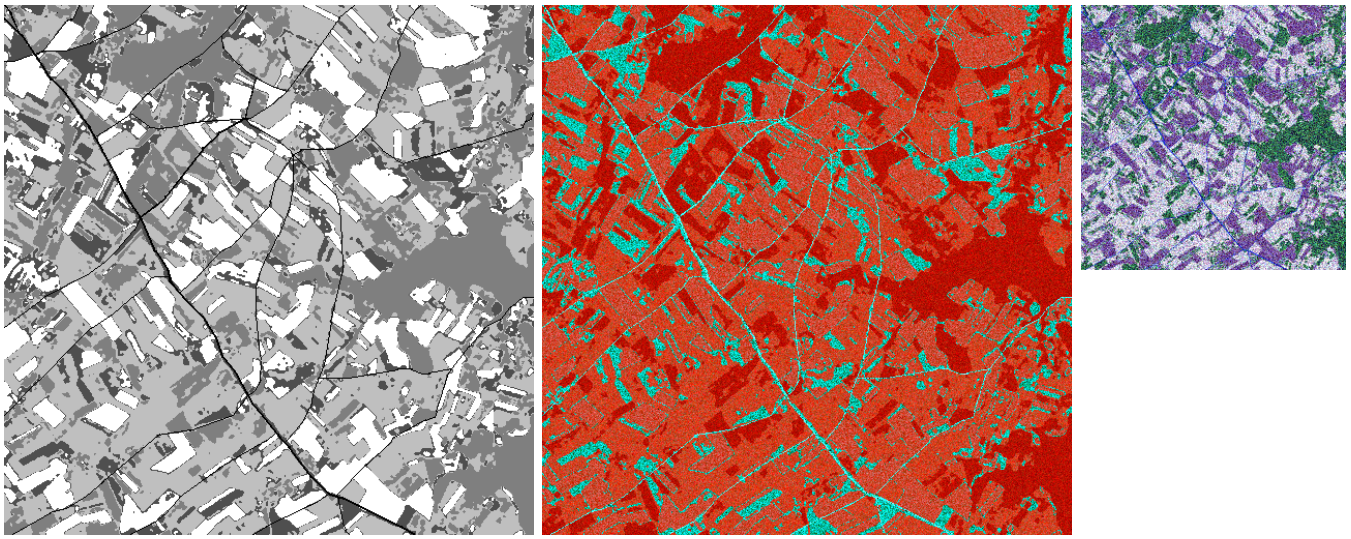


Fig. 1. Illustration of the multi-scale model in the case of two resolutions.

We also assume that the pixel dimensions at lower resolutions are entire multiples of the pixel dimensions at the reference resolution (level 1), and that the images are perfectly overlapping. An extension to partly overlapping pixels and images is possible, but it would imply a considerably more complicated notation and implementation.



(a) (b) (c)

Fig. 2. Simulated images for multi-scale classification: (a) 512×512 class label image with 5 classes. (b) 3-band 512×512 simulated image with class characteristics derived from SPOT XS data. (c) 6-band 256×256 simulated image with class characteristics derived from Landsat TM data (3 bands shown).

In order to specify $p(\mathbf{y}^j|\mathbf{z})$, we introduce the notation:

$s_j(v)$ =The set of pixels at the reference resolution which is contained in pixel v at resolution r_j.

m_j =The number of pixels in $s_j(v)$.

$v_j(i)$ =The pixel at resolution r_j containing pixel i at the reference resolution.

Referring to Fig. 1, the data at coarser resolutions are modeled as mixed pixels, i.e., we introduce hidden variables $\widetilde{\mathbf{y}}_i^j$ at the reference level that sums up to the observed pixel values $\mathbf{y}_v^j$ at the coarser resolution level j:

$$\mathbf{y}_v^j = \frac{1}{m_j} \sum_{i \in s_j(v)} \widetilde{\mathbf{y}}_i^j \tag{2}$$

We can consider $\widetilde{\mathbf{y}}_i^j$ as the observation that *would* be obtained if a sensor existed that had the radiometric properties of the sensor at level j, but acquired images at the reference resolution. A similar assumption has been used in [5] and [6]. We further assume

$$\widetilde{\mathbf{y}}_i^j|\mathbf{z} \sim N(\boldsymbol{\mu}_{z_i}^j, \boldsymbol{\Sigma}_{z_i}^j) \tag{3}$$

where $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the Normal distribution with expectation vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.

III. Classification

Assuming all parameters involved are known, classification is based on the posterior distribution

$$p(\mathbf{z}|\mathbf{y}; \boldsymbol{\theta}) \propto p(\mathbf{z}) \prod_{j=1}^{p} p(\mathbf{y}^j|\mathbf{z}). \tag{4}$$

The maximum *a posteriori* (MAP) solution is obtained by maximizing (4) with respect to the complete vector $\mathbf{z}$. Explicit MAP solutions have been obtained in [7] for problems with all observations given at the same resolution. Although such algorithms (based on integer programming and Lagrangian based methods) should be possible to develop also for the problem at hand, we here consider a simpler version based on the ICM algorithm [4].

The ICM algorithm consists of sequential optimization of the components of $\mathbf{z}$ by maximizing $p(z_i|z_{i'}, i' \neq i, \mathbf{y})$. Using (1) and (4)

$$p(z_i|z_{i'}, i' \neq i, \mathbf{y}) \propto \exp\{\alpha_{z_i} + \beta \sum_{i' \sim i} I(z_i = z_{i'})\} \prod_{j=1}^{p} p(\mathbf{y}_{v_j(i)}^j|\mathbf{z}). \tag{5}$$

For each ICM iteration and for each pixel i, (5) is calculated for $z_i = 1, \ldots, K$, and the class with the highest probability is retained. From (2) and (3), we get

$$\mathbf{y}_v^j|\mathbf{z} \sim N(\frac{1}{m_j} \sum_{i \in s_j(v)} \boldsymbol{\mu}_{z_i}^j, \frac{1}{m_j} \sum_{i \in s_j(v)} \boldsymbol{\Sigma}_{z_i}^j) \tag{6}$$

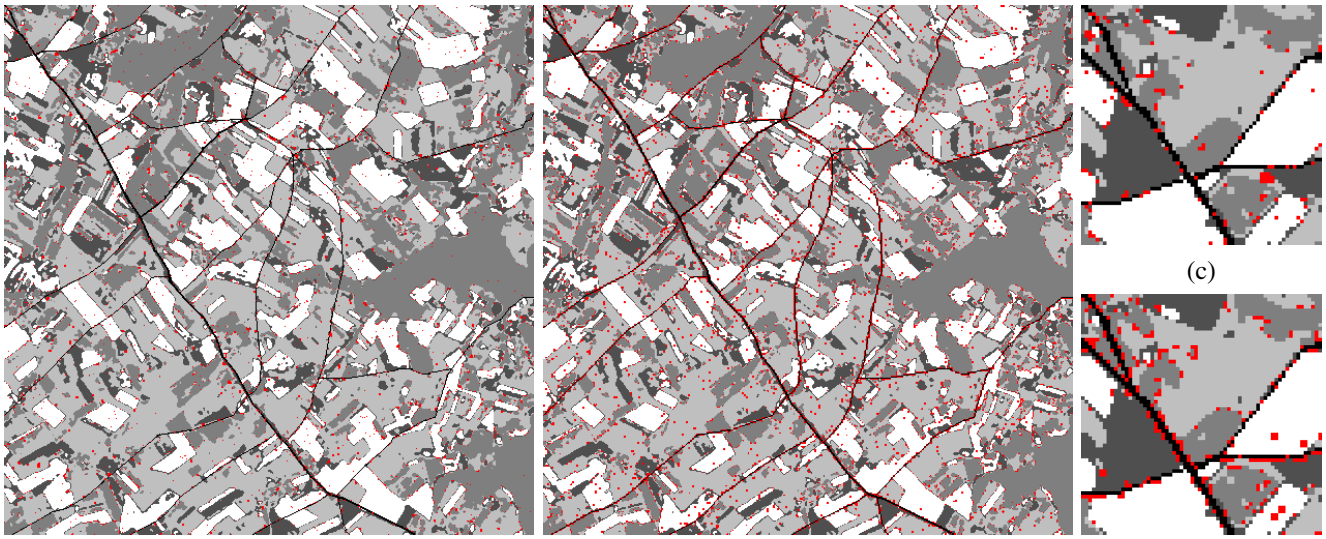
making (5) relatively easy to compute.

IV. Parameter Estimation

The model described in section II contains many parameters that need to be estimated or specified by other means. Due to the presence of data at several resolutions, ML estimates are computationally expensive. We have therefore applied a simpler approach based on [4]. The details are given in [2].

V. Results

The proposed multi-scale classification scheme has been tested on a simulated data set, where the class properties are derived from real SPOT XS and Landsat TM images. We used simulated images because the new method only can be expected to bring improvements in the presence of fine structures such as roads and transitions corresponding to region boundaries, whereas ground truth for real images generally is collected well inside homogeneous regions.



(a) (b) (c) (d)

Fig. 3. Classification result of (a) the multi-scale classification scheme and (b) the corresponding single-scale classification scheme applied after resampling of the second image to the reference resolution. Correctly classified pixels are shown in gray levels and misclassified pixels are shown in red. The upper left corners are magnified to show some representative details of (c) the multi-scale result and (d) the single-scale result.

The class label image that we used is shown in Fig. 2 (a). It represents an agricultural scene and is derived from a class image used in a previous study [8], by manually adding roads, so that the number of classes becomes 5. It is not a realization of a MRF. The 3-band image shown in Fig. 2 (b) was obtained by simulating multivariate Gaussian noise according to the class properties estimated from a real SPOT XS image, and by attributing the signal vectors to the pixels of the ideal class image. Likewise, a 6-band image was simulated from the class properties of a real Landsat TM image. However, it was done according to the multi-scale model to yield an image that has 2×2 times lower resolution than the initial class image, and mixed pixels at region boundaries. A square of 100×100 pixels in the ideal class image was defined as training set, and the rest of the image constituted the test set.

The result of the proposed multi-scale classification scheme is shown in Fig. 3 (a). To compare the multi-scale approach with single-scale analysis of the data set, we resampled the image with Landsat TM properties to the reference resolution using nearest neighbor interpolation, and concatenated the resulting image bands with those of the image with SPOT XS properties, so that we got a 512×512 image with 9 bands. The ICM loops for parameter estimation and classification were carried out, the only difference from the multi-scale algorithm being that the computations related specifically to the multi-scale model are let out. The result is shown in Fig. 3 (b). Magnified extracts of the classification results are presented in Fig. 3 (c) and (d). Single-scale analysis was also performed separately on the image with SPOT XS properties and the resampled image with Landsat TM properties.

The most important performance measures are summarized in Table I. We see that there are only 9.1 % misclassified pixels when applying the single-scale ICM scheme to the image with SPOT XS properties only. The same method applied to the concatenation of this image and the resampled version of the image with Landsat TM properties further reduces the proportion of misclassified pixels by 4.3 %, and another 2.6 % are eliminated with the new multi-scale method, so that only 2.2 % of misclassified pixels remain. However, the last improvement is obtained at the cost of 5 times higher computing time.

TABLE I

PERFORMANCE SUMMARY

Method	Data	Accuracy	Time
Multi-scale ICM	XS + TM	97.8 %	8 m 33 s
Single-scale ICM	XS + TM	95.2 %	1 m 42 s
Single-scale ICM	XS	90.9 %	45 s
Single-scale ICM	TM	78.3 %	54 s

As can be seen from the classification results in Fig. 3, the new multi-scale approach significantly improves the classification accuracy near fine structures and regions boundaries.

REFERENCES

[1] J. Besag, "Towards Bayesian image analysis," *Journal of Applied Statistics*, vol. 16, no. 3, pp. 395–407, 1989.

[2] G. Storvik, R. Fjørtoft, and A. H. S. Solberg, "Parameter estimation and classification of multi-scale remote sensing data," Norwegian Computing Center (NR), Oslo, Norway, Internal report SAMBA/06/2003, May 2003.

[3] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 721–741, 1984.

[4] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society, Series B*, vol. 48, no. 3, pp. 259–302, 1986.

[5] J. C. Price, "Combining multispectral data of differing spatial resolution," *IEEE Trans. Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1199–1203, 1999.

[6] B. Zhukov, D. Oertel, F. Lanzl, and R. G., "Unmixing-based multisensor multiresolution image fusion," *IEEE Trans. Geoscience and Remote Sensing*, vol. 37, no. 3, pp. 1212–1226, 1999.

[7] G. Storvik and G. Dahl, "Lagrangian based methods for finding MAP solutions for MRF models," *IEEE Trans. Image Processing*, vol. 9, no. 3, pp. 469–479, 2000.

[8] R. Fjørtoft, Y. Delignon, W. Pieczynski, M. Sigelle, and F. Tupin, "Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields," *IEEE Trans. Geoscience and Remote Sensing*, vol. 41, no. 3, 2003.