Fusion of Earth Data, Sophia Antipolis, France, 28-30 January 1998

Fusion of SPOT HRV XS and Orthophoto Data Using a Markov Random Field Model

Bjarne K. Ersbøll, Knut Conradsen, Rasmus Larsen, Allan A. Nielsen and Thomas H. Nielsen
Department of Mathematical Modelling, Technical University of Denmark
Building 321, DK-2800 Lyngby, Denmark
Phone +45 4588 1433, Direct +45 4525 3413, Fax +45 4588 1397
E-mail be@imm.dtu.dk, Internet www.imm.dtu.dk

ABSTRACT: In remote sensing there is usually a trade-off between spatial resolution and spectral resolution. It is often desirable to present a combination of the data which merges the low resolution spectral (colour) information with the high resolution spatial information. If the ratio in spatial resolution between the greyscale image and the colour image becomes too great the image resulting from usual itensity substitution techniques will tend to look as if it has “squares” of colour information overlaid on the greyscale image. A new way of “fusing” the colour information with the greyscale information is presented here. The technique is based on Markov random field (MRF) assumptions and the maximum a posteriori (MAP) estimate is found by means of iterated conditional modes (ICM). The design of the specific Markov random field used here allows the user to control three properties of the resulting “merged” image: 1) degree of colour smoothness between neighbouring pixels, 2) degree of confidence in the original colour information (e.g. SPOT HRV XS), and 3) degree of confidence in the original panchromatic information (e.g. orthophoto). As a possible fourth control element, we may exclude neighbours belonging to other segments of the segmented orthophoto from the neighbourhood of the MRF. The data used here as an example are from the town of Sorø in Denmark. The data consist of a digital orthophoto (62.5 cm pixels) and a SPOT HRV XS (20 m pixels) giving a linear ratio of 32:1 between orthophoto and SPOT pixels. In this presentation, we compare results based on neighbourhoods with and without the exclusion of neighbours in other segments.

1 INTRODUCTION

In remote sensing one is often given the option of choosing between two different types of data: high (spatial) resolution greyscale imagery on one hand and low (spatial) resolution colour imagery on the other.

One example of such data is SPOT HRV P (10 m pixels) and SPOT HRV XS (20 m pixels). Another more extreme example is panchromatic orthophoto (62.5 cm pixels) and SPOT HRV XS (20 m pixels).

It is often desirable to give some sort of presentation which combines the high spatial resolution of the greyscale image and the colour information from the colour (multispectral) image.

An often used way to do this consists of transforming the colour (RGB) image to Munsell (intensity-hue-saturation or IHS) coordinates, replacing the intensity with the high resolution greyscale image and transforming back to RGB.

However, if the ratio in spatial resolution between the greyscale image and the colour image becomes too great (e.g. 1 m pixels and 30 m pixels) the resulting image will tend to look as if it has “squares” of colour information overlaid on the greyscale image. This can to some extent be avoided by smoothing the colour information prior to the transformation.

A new way of “fusing” the colour information with the greyscale information is given here. The technique is based on Markov random field (MRF) assumptions and the maximum a posteriori (MAP) estimate is found by using the iterated conditional modes (ICM) algorithm (Besag, 1986).

2 DATA

The data (kindly provided and geocoded by the National Survey and Cadastre of Denmark) are

- a digitised panchromatic orthophoto over the Danish town of Sorø. The pixels are 62.5 cm $\times$ 62.5 cm. The full scene is 7200 $\times$ 7200 pixels, the subscene chosen for the example is 512 $\times$ 512 pixels.
- a SPOT HRV XS image of the same area. The pixels are 20 m $\times$ 20 m. This corresponds to 225 $\times$ 225 pixels for the full orthophoto and 16 $\times$ 16 pixels for the subscene.

3 METHODS

The Markov random field (MRF) assumption states that the probability of a pixel having a particular value depends only on the pixels in its neighbourhood, N_i

$$P(X_i = x_i | x_s, s \neq i) = P(X_i = x_i | x_s \in N_i, s \neq i),$$

where X_i in this case is a three dimensional random variable with scalar values for XS1, XS2 and XS3 respectively.
Assume there is a true but unknown high resolution three dimensional (i.e. X1-X3) image X. Assume that we observe the image Y consisting of the low resolution XS1-XS3 data and the high resolution orthophoto data. Bayes formula then gives us

$$P(X = x | Y = y) = \frac{P(Y = y | X = x) P(X = x)}{P(Y = y)}.$$

The task is to find $X = x$ which maximises this globally. This is practically impossible. However, Hammersley-Clifford's theorem states an equivalence between the local MRF description (via so-called potentials U_i) and the global description (via the so-called energy $U(x)$) by means of a Gibbs-distribution (the prior)

$$P(X = x) = \frac{1}{Z} \exp(\sum_{N_i} -U_i) = \frac{1}{Z} \exp(-U(x)).$$

Bayes formula can now be expressed as

$$P(X = x | Y = y) = \frac{1}{Z} \exp(-U_{\text{prior}} - U_{\text{observed}}).$$

The maximum can now be found iteratively by simulated annealing (Geman and Geman, 1984) or iterated conditional modes (Besag, 1986).
We use three terms in the energy function

1. *prior assumption*: that the image appears smooth, i.e.
   $U_{\text{prior}}(x_i) = \alpha \sum_{j \in N_i} ||x_i - x_j||^2$,
2. *observed information concerning colours*: the resulting image should honour the SPOT HRV XS image, i.e.
   $U_{\text{obs,RGB}}(y_{i,\text{SPOT}} | x_i) = \beta ||x_i - y_{i,\text{SPOT}}||^2$, and

3. *observed information concerning intensity*: the resulting image should honour the orthophoto, i.e.
   $U_{\text{obs,ortho}}(y_{i,\text{ortho}}|x_i) = \gamma(\frac{1}{3}(x_{i,R} + x_{i,G} + x_{i,B}) - y_{i,\text{ortho}})^2$.

The above optimisation is carried out with two different neighbourhood models, namely 1) the ordinary four nearest neighbours, and 2) those of the four nearest neighbours that are positioned in the same segment of the orthophoto as the running pixel.
The segmentation of the orthophoto is carried out by subtracting the image itself from a greyscale eroded version followed by a partitioning into watersheds.

4 RESULTS

Figure 1 shows the original SPOT HRV XS and orthophoto data. Figures 2 and 3 show the fused version without and with exclusion of neighbours not positioned in the same segment of the orthophoto as the centre pixel. Since the objective function is convex in this case we use ICM for the global maximisation.

5 CONCLUSION

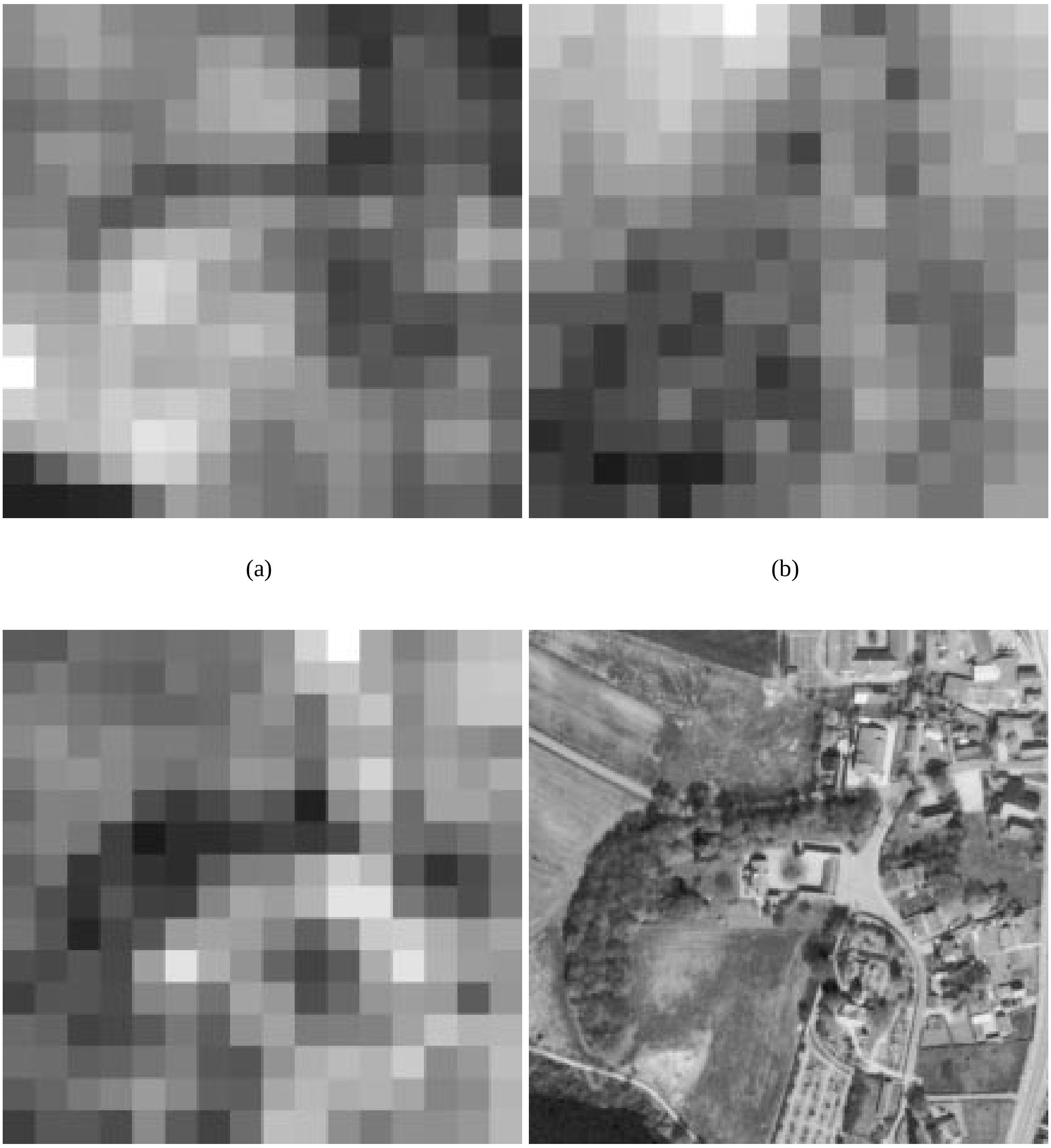
Using a Markov random field description of the image it was possible to construct an image-merging algorithm giving control over the degree of confidence in the prior information (α, smoothness in colour) and the observed information (β, confidence in SPOT HRV XS and γ, confidence in orthophoto).
A clear difference between the segmentation and no-segmentation situations is seen. With the segmentation a more distinct and "edgy" visual appearance is obtained. Note, that in the segmentation situation colours tend not to smear across bounderies. This is clearly seen along roads and on roof tops. However, the result relies heavily on the segmentation of the orthophoto which may not be of the same quality in all parts of the image. Particularly the balance between "over-" and "under-segmentation" in different parts of the image is a problem. The technique as presented here cannot be considered fully developed.

ACKNOWLEDGEMENTS

Computer code for the erosion was written by Dr J. Michael Carstensen, IMM. Computer code for the watershed algorithm was written by Dr P. Ulrik Skands, IMM. The geocoded data were kindly provided by KMS, Kort- og Matrikelstyrelsen (the National Survey and Cadastre of Denmark).

REFERENCES

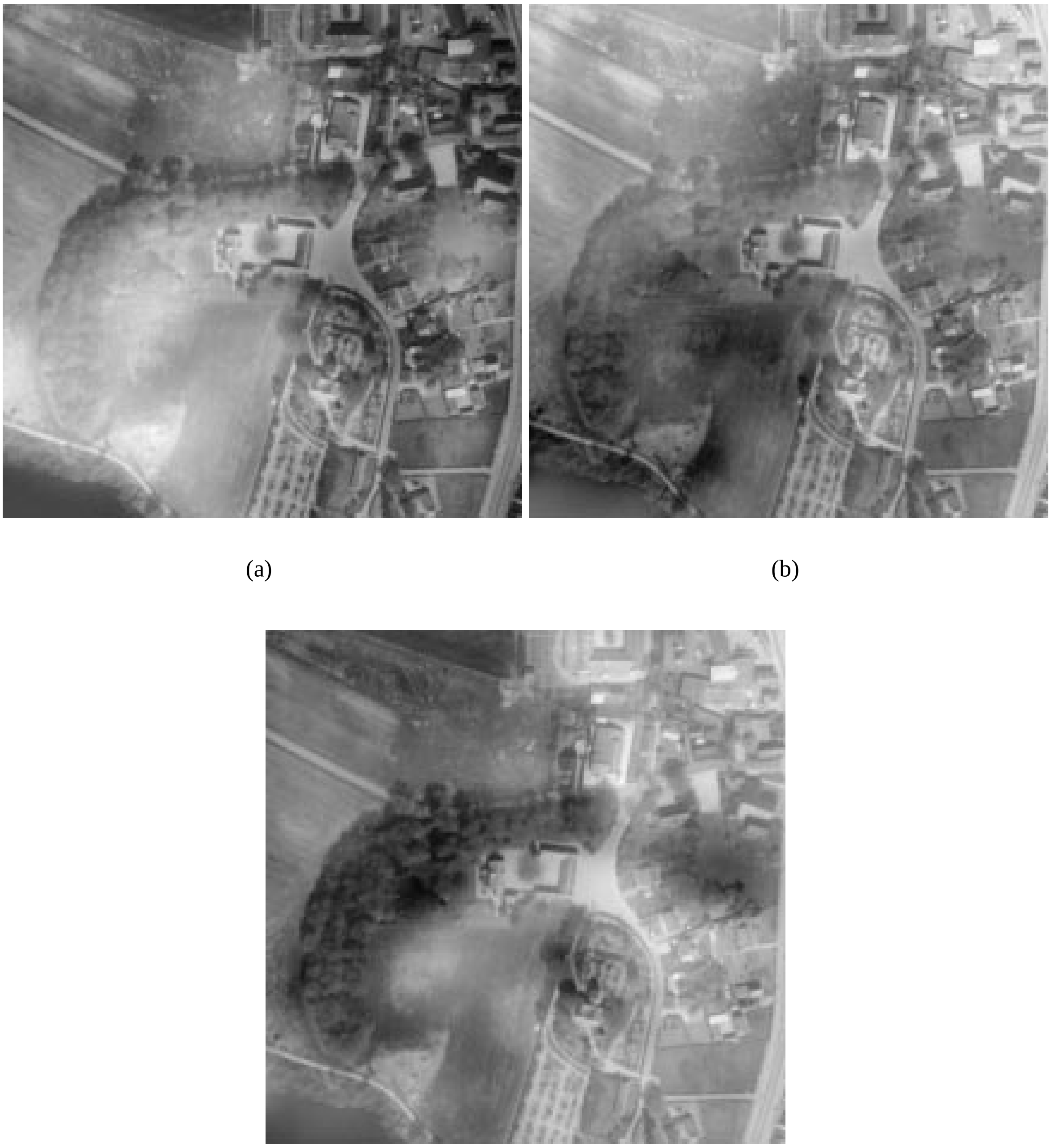
Besag, J. (1986). On the Analysis of Dirty Pictures. *Journal of the Royal Statistical Society, Series B*, **48**, pp. 259-302.
Geman, S. and Geman, D. (1984). Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **6**(6), pp. 721-741.
Nielsen, T.H. (1994). *Fusion af Ortofoto og Farver*. In Danish. Department of Mathematical Modelling, Master Thesis No. 1994-35, Technical University of Denmark. ISSN 0909-6256. 188 p.



(a) (b)

(c) (d)

Figure 1: The original data, (a) SPOT HRV XS1 (b) SPOT HRV XS2 (c) SPOT HRV XS3 (d) orthophoto



(a) (b) (c)

Figure 2: The fused SPOT HRV XS and orthophoto data, all four nearest neighbours are used, (a) XS1 (b) XS2 (c) XS3

(a) (b)

(c)

Figure 3: The fused SPOT HRV XS and orthophoto data, those of four nearest neighbours that are in same segment of orthophoto as center pixel are used, (a) XS1 (b) XS2 (c) XS3