# Learning Spatial Filters for Multispectral Image Segmentation

Devis Tuia*, Gustavo Camps-Valls*,
**Rémi Flamary****, Alain Rakotomamonjy**

*Image Processing Laboratory (IPL)
Universitat de València, Spain

**LITIS EA 4108, Université de Rouen
76800 Saint Etienne du Rouvray, France

August 24, 2010

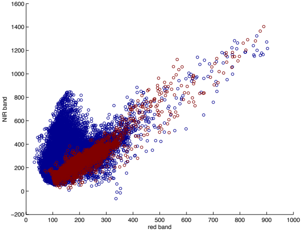
# Multispectral Image segmentation

- In multispectral images we have high spatial variability of the spectral signature
- VHR images allows us better recognition, but noisy maps
- Strong intraclass variance, higher than interclass
- Including spatial, not only spectral info, is mandatory!

  ⇒ *Spatial regularization*

image

ground thruth

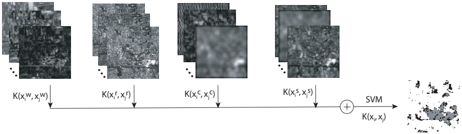



Class "Buildings' (red) against the others (blue)'

# Spatial Filtering

Spatial filtering solves such problems:

- Mathematical morphology [Benediktsson et al., 2005, Tuia et al., 2010]
- Geometrical features [Inglada, 2007]
- Composite kernels with spatial filters [Camps-Valls et al., 2006]




But remain the problem of defining

- what kind of features,
- at what scale, ...

In this paper we propose to
learn the spatial filter that maximizes separability of the classes
in a SVM framework.

# Definitions

- $\mathbf{X} \in \mathbb{R}^{r_1 \times r_2 \times d}$ is an image containing $r_1 \times r_2$ pixels $\in \mathbb{R}^d$.
- $\mathbf{X}_{i,j,k} = \mathbf{X}_{\mathbf{p},k}$ ais the $k$th component of pixel $\mathbf{p} = (i,j)$.
- 2D convolution filter band-by-band:

$$\widetilde{\mathbf{X}}_{\mathbf{p},k} = \sum_{u=1,v=1}^{f,f} \mathbf{F}_{u,v,k} \; \mathbf{X}_{\mathbf{p}+(u,v)-f_0,k}$$

  where $f_0 = f/2$ and $\mathbf{F} \in \mathbb{R}^{f \times f \times d}$.
- RBF kernel between filtered pixels:

$$\widetilde{K}_{\mathbf{p},\mathbf{q}} = k(\widetilde{\mathbf{X}}_{\mathbf{p},.}, \widetilde{\mathbf{X}}_{\mathbf{q},.}) = \exp\left(-\frac{||\widetilde{\mathbf{X}}_{\mathbf{p},.} - \widetilde{\mathbf{X}}_{\mathbf{q},.}||^2}{2\sigma^2}\right), \qquad (1)$$

  where $\sigma$ is the kernel width or bandwidth.

# Optimization Problem

## Large Margin Filtering [Flamary et al., 2010]

$$\min_{g,\mathbf{F}} \Bigg\{ \underbrace{\frac{1}{2}\|g\|^2 + \frac{C}{n}\sum_{\mathbf{p}\in\mathcal{S}_l} H(\mathbf{Y_p}, g(\widetilde{\mathbf{X}}_{\mathbf{p},.}))}_{\text{SVM objective function}} + \underbrace{\lambda\Omega(\mathbf{F})}_{\text{Filter regularization}} \Bigg\} \quad (2)$$

where:

- $H(\mathbf{Y_p}, g(\widetilde{\mathbf{X}}_{\mathbf{p},.}) = \max(0, 1 - \mathbf{Y_p} \cdot g(\widetilde{\mathbf{X}}_{\mathbf{p},.}))$ is the SVM hinge loss.
- $C$ and $\lambda$ are the regularization parameters.
- $\Omega(\cdot)$ is a 3D Frobenius Norm: $\Omega(\mathbf{F}) = \sum_{u,v,k}^{f,f,d} \mathbf{F}_{u,v,k}^2$
- $g(\cdot)$ is the SVM decision function:

$$g(\widetilde{\mathbf{X}}_{\mathbf{p},.}) = \sum_{\mathbf{q}\in\mathcal{S}_l} \alpha_\mathbf{q} \mathbf{Y_q} k(\widetilde{\mathbf{X}}_{\mathbf{q},.}, \widetilde{\mathbf{X}}_{\mathbf{p},.}) + b, \quad (3)$$

where $\alpha_\mathbf{p}$ are the dual variables of problem.

# Solving the problem

## Approach

- Convex problem for a fixed **F**.
- We can always find the optimal decision function $g^*$ for a fixed **F**.
- Do gradient descent on **F** and stay in the optimal $g^*$ space [Bonnans and Shapiro, 1998]:

$$\min_{\mathbf{F}} J(\mathbf{F}) = \min_{\mathbf{F}} J'(\mathbf{F}) + \lambda\Omega(\mathbf{F}) \tag{4}$$

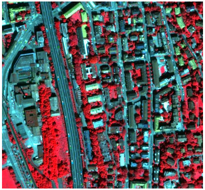
with:

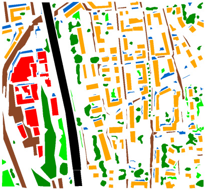
$$J'(\mathbf{F}) = \min_{g} \left\{ \frac{1}{2}\|g\|^2 + \frac{C}{n} \sum_{\mathbf{p}\in\mathcal{S}_l} H(\mathbf{Y}_{\mathbf{p}}, g(\widetilde{\mathbf{X}}_{\mathbf{p},.})) \right\} \tag{5}$$

## Algorithm [Flamary et al., 2010]

- Conjugate Gradient descent on **F** + linesearch.
- Solve a SVM at each cost calculation in the linesearch.

# Dataset and experimental setup





## Dataset

- VHR QuickBird image of the city of Zurich, Switzerland.
- 7 classes, difficult to discriminate 'buildings' classes ('residential' vs 'commercial'). If merged, difficulty to discriminate 'buildings' and 'roads'

## Compared Models

1. SVM pixel classifier.
2. AvgSVM, averaged pixel classifier.
3. WinSVM, classification of a window of pixels.
4. KF-SVM, Large margin filtering.

# Binary Classification

| Method | Class | Training Pixels | #Class Pixels | AUC | Kappa |
|---|---|---|---|---|---|
| SVM | Residential | $\sim$ 5000 | 2000 | 0.904 | 0.638 |
| AvgSVM | | | | 0.916 | 0.689 |
| WinSVM | | | | **0.947** | 0.730 |
| KF-SVM | | | | 0.938 | **0.742** |
| SVM | Buildings* | $\sim$ 4000 | 1000 | 0.938 | 0.706 |
| AvgSVM | | | | 0.946 | 0.779 |
| WinSVM | | | | 0.970 | 0.807 |
| KF-SVM | | | | **0.974** | **0.815** |

* Pixels from classes ‘Residential’ and ‘Commercial’.

## Results

- The estimated Area Under the ROC Curve (AUC) and Kappa coefficient are computed.
- Improving over the SVM classification and average Filtering.
- Similar results between KF-SVM and WinSVM (slightly better Kappa).

# Multiclass classification

| Method | Classes | Filter size | Training Pixels | [%]OA | Kappa |
|---|---|---|---|---|---|
| SVM | | | | 75.11 | 0.685 |
| AvgSVM | 7 | 9 | $\sim 5000$ | 83.68 | 0.796 |
| WinSVM | | | | 82.98 | 0.785 |
| KF-SVM | | | | **85.32** | **0.816** |
| SVM | | | | 83.04 | 0.772 |
| AvgSVM | 6* | 9 | $\sim 5000$ | 89.48 | 0.860 |
| WinSVM | | | | **91.71** | **0.889** |
| KF-SVM | | | | 91.45 | 0.885 |

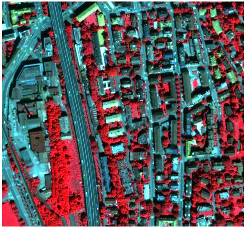
* Pixels from classes ‘Residential’ and ‘Commercial’.

## Results

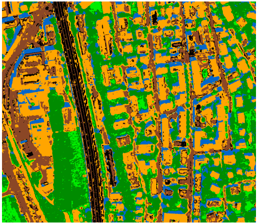
- WinSVM and KF-SVM give similar results and both outperform SVM and AvgSVM.
- But with KF-SVM, only pixels are classified.
- Optimal preprocessing done by filtering.

# Segmentation Visualization

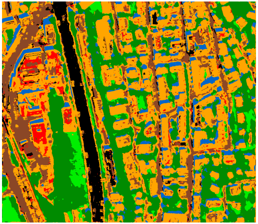
Image

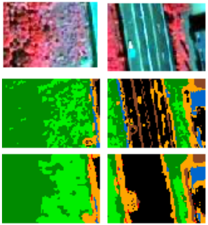


SVM



KF-SVM





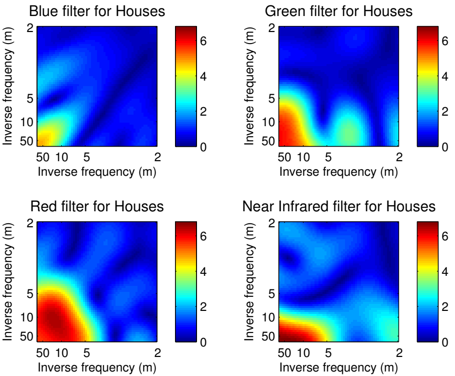
## Results

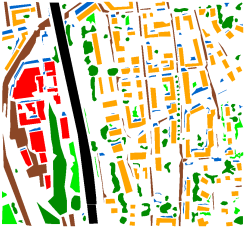
- Smoothed borders thanks to the filtering.
- 'Commercial Buildings' (red class) detected over 'Buildings' (orange class).
- What about the filters ? $\Rightarrow$ Fourier transform

# Filter Visualization (1)

## Class: Houses, Residential buildings

Magnitude of the FT for different components



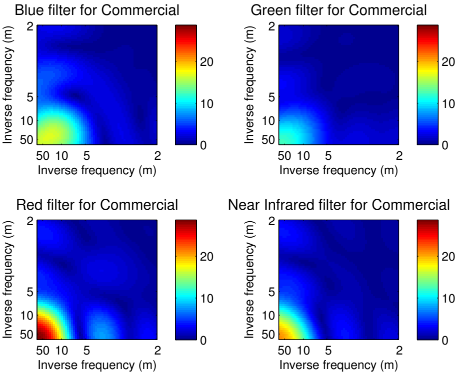
(Orange)

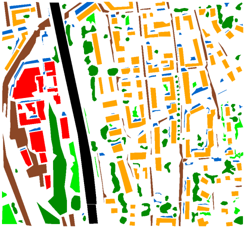
- Low pass but larger band (houses are small).
- Green, Red and InfraRed are selected.

# Filter Visualization (2)

## Class: Commercial buildings

Magnitude of the FT for different components



(Red)

- Low pass but small band to detect large buildings.
- Red is the most discriminant feature.

# Conclusion

## Large Margin Filtering

- Method to learn jointly a SVM pixel classifier and a spatial filtering.
- Large margin spatial filtering/Preprocessing.
- Possibility to use other classifier after filtering, e.g. GMM.
- Visualization for the filtering, no black box approach.

## Future Work

- Propose other regularization terms.
- Going local, a global filter is limited.
- Test the method in hyperspectral images, where stacking approaches fail.

# References


[Benediktsson et al., 2005] Benediktsson, J., Palmason, J. A., and Sveinsson, J. R. (2005).
Classification of hyperspectral data from urban areas based on extended morphological profiles.
*IEEE Trans. Geosci. Remote Sens.*, 43(3):480–490.

[Bonnans and Shapiro, 1998] Bonnans, J. and Shapiro, A. (1998).
Optimization problems with pertubation : A guided tour.
*SIAM Review*, 40(2):202–227.

[Camps-Valls et al., 2006] Camps-Valls, G., Gómez-Chova, L., Muñoz-Marí, J., Vila-Francés, J., and Calpe-Maravilla, J. (2006).
Composite kernels for hyperspectral image classification.
*IEEE Geosci. Remote Sens. Lett.*, 3(1):93–97.

[Flamary et al., 2010] Flamary, R., Labbé, B., and Rakotomamonjy, A. (2010).
Large margin filtering for signal sequence labeling.
In *IEEE Conference on Acoustic, Speech and Signal Processing ICASSP*, Austin, Texas, USA.

[Inglada, 2007] Inglada, J. (2007).
Automatic recognition of man-made objects in high resolution optical remote sensing images by SVM classification of geometric image features.
*ISPRS Journal of Photogrammetry Rem. Sens.*, 62:236–248.

[Tuia et al., 2010] Tuia, D., Ratle, F., Pozdnoukhov, A., and Camps-Valls, G. (2010).
Multi-source composite kernels for urban image classification.
*IEEE Geosci. Remote Sens. Lett., Special Issue ESA_EUSC*, 7(1):88–92.