Interactive Classification Oriented Superresolution of Multispectral Images

P. Ruiz

Department of Computer Science and Artificial Intelligence, Universidad de Granada
Granada, 18071, España
E-mail: mataran@decsai.ugr.es
decsai.ugr.es/vip

J. V. Talents

Image Processing Laboratory (IPL), Universitat de Valencia
Valencia, 46071 España
E-mail: juatano@.uv.es

J. Mateos and R. Molina

Department of Computer Science and Artificial Intelligence, Universidad de Granada
Granada, 18071, España
E-mail: {rms, jmd}@decsai.ugr.es

A.K. Katsaggelos

Department of Electrical Engineering and Computer Science, Northwestern University
Evanston, IL, 60208-3118 USA
E-mail: aggk@eecs.northwestern.edu

Classification techniques are routinely utilized on satellite images. Pansharpening techniques can be used to provide super resolved multispectral images that can improve the performance of classification methods. So far, these pansharpening methods have been explored only as a preprocessing step. In this work we address the problem of adaptively modifying the pansharpening method in order to improve the precision and recall figures of merit of the classification of a given class without significantly deteriorating the performance of the classifier over the other classes. The validity of the proposed technique is demonstrated using a real Quickbird image.

Keywords: Pansharpening, super-resolution, classification, LDA, SVM.

1. Introduction

Satellite images are of great interest due to the numerous applications they can be utilized. Drawing maps, delimitation of parcels, studies on hydrology, forest or agriculture are just a few examples where these images are used.

Due to physical and technological constraints, satellites usually have sensors that capture two types of images. One sensor captures a multispectral (MS) image composed of several spectral bands with low spatial resolution (LR). The other sensor captures a high spatial resolution (HR) image, named panchromatic (PAN) image, with a low spectral resolution. While the first image allows to distinguish features spectrally but not spatially, the second allows to distinguish features spatially but not spectrally.

Pansharpening is an image fusion approach that combines the LR MS and PAN images to obtain an image with the spectral resolution of the MS image and the spatial resolution of the PAN image. Many techniques have been proposed in the literature to carry out the pansharpening procedure (see Ref. 1 for a complete review of pansharpening methods).

Many satellite image applications involve the classification of pixels in an image into a number of classes. In supervised classification, starting from a small set of samples previously labeled by the user, classification is carried out automatically by the classifiers. Bruzzone *et al.*[2] showed that the use of pansharpening methods that do not introduce significant spectral distortion helps the classifier to obtain higher accuracy, especially for pixels at the borders of objects.

While in the past pansharpening techniques have only been used as a preprocessing step, in this work we address the problem of adaptively modifying the pansharpening method in order to improve the precision and recall figures of merit of the classification of a given class without deteriorating the performance of the classifier over the other classes.

The rest of paper is organized as follows: In section 2 we describe the pansharpening technique we use. The used classifiers are briefly explained in section 3. The proposed method to estimate the pansharpening parameters to improve the performance of the classifier on a given class is described in section 4. Section 5 presents experimental results on real data. Finally, section 6 concludes the paper.

2. Pansharpening Algorithm

In this paper we use the pansharpening method proposed by Amro *et al.*[3] and the parameter estimation procedure described in Ref. 1. This method

makes use of the non-subsampled contourlet transform[4] (NSCT) to decompose the details of the PAN and each band of the MS image into different scales and different directions. Then, the hierarchical Bayesian framework is used to model those observations and their relations with the original high resolution multispectral image and Bayesian inference is applied to estimate the HR MS image and the model parameters. Let us now explain in detail the used pansharpening method.

The used contourlet based pansharpening algorithm takes as input the PAN image, x, of size $p = m \times n$, and the observed LR MS image, Y, with B bands, $Y_b, b = 1, \ldots, B$, each of size $P = M \times N$ pixels with $M < m$ and $N < n$. Initially, each band of the LR MS image Y is upsampled to the size of the PAN image by bicubic interpolation. We will denote by s_b each band b of the $p = m \times n$ upsampled image.

Then, using the NTSC transform we can write the PAN and the upsampled MS images as:

$$x = x^r + \sum_{l=1}^{L}\sum_{d=1}^{D} x^{ld}, \quad s_b = s_b^r + \sum_{l=1}^{L}\sum_{d=1}^{D} s_b^{ld}, \quad b = 1, \ldots, B \tag{1}$$

where the superscript r denotes the residual (low pass filtered version) NSCT coefficient band and the superscript ld refers to the detail bands, with $l = 1, \ldots, L$, representing the scale and $d = 1, \ldots, D$, representing the direction for each coefficient band. The pansharpening goal is to estimate the HR MS image coefficients y_b^{ld} from the observed x^{ld} and s_b^{ld} coefficients. Finally, each band of the pansharpened HR MS image will be obtained by the inverse NSCT from the corresponding residual band of the upsampled MS image s_b^r and the estimated detail bands y_b^{ld}.

We will model the coefficient bands using the hierarchical Bayesian framework. This framework has two stages. In the first stage, knowledge about the structural form of the noise in the coefficients bands and the structural behavior of the HR MS image coefficients is used in forming $p(s_b^{ld}, x^{ld}|y_b^{ld}, \Omega_b^{ld})$ and $p(y_b^{ld}|\Omega_b^{ld})$, respectively. These noise and image models depend on the unknown parameters Ω_b^{ld} that need to be estimated. In the second stage a hyperprior on the parameters is defined, thus allowing the incorporation of information about these hyperparameters into the process. Let us define the probability distribution involved in each stage.

Following Refs. 3,5, we chose a prior model based on the Total Variation (TV) for the HR MS image coefficient bands, y_b^{ld}, given by

$$p(y_b^{ld}|\alpha_b^{ld}) \propto (\alpha_b^{ld})^{p/2} \exp\left\{-\alpha_b^{ld} TV(y_b^{ld})\right\}, \tag{2}$$

with $TV(y_b^{ld}) = \sum_{i=1}^{p} \sqrt{(\Delta_i^h(y_b^{ld}))^2 + (\Delta_i^v(y_b^{ld}))^2}$ where $\Delta_i^h(y_b^{ld})$ and $\Delta_i^v(y_b^{ld})$ represent the horizontal and vertical first order differences at pixel i, respectively, and α_b^{ld} is the model parameter of the MS band b coefficients at level l and direction d. The idea behind this model is to consider the coefficient bands as a set of relatively smooth regions separated by strong edges, such as the coefficients of the NSCT.

Since the MS bands coefficients and the PAN image coefficients are independent given the HR MS image coefficients, we define $p(s_b^{ld}, x^{ld}|y_b^{ld}, \Omega_b^{ld}) = p(s_b^{ld}|y_b^{ld}, \Omega_b^{ld}) \times p(x^{ld}|y_b^{ld}, \Omega_b^{ld})$. The conditional distribution of the upsampled MS coefficients given the HR MS coefficients is defined as[3]

$$p(s_b^{ld}|y_b^{ld}, \beta_b^{ld}) \propto (\beta_b^{ld})^{p/2} \exp\left\{-\frac{1}{2}\beta_b^{ld} \left\|s_b^{ld} - y_b^{ld}\right\|^2\right\}, \tag{3}$$

where β_b^{ld} is the inverse of the unknown noise variance of the detail band at level l and direction d of the MS band b. The relationship between the HRMS band coefficients and the PAN image is modeled by the conditional probability distribution

$$p(x^{ld}|y_b^{ld}, \gamma_b^{ld}) \propto (\gamma_b^{ld})^{p/2} \exp\left\{-\frac{1}{2}\gamma_b^{ld} \left\|x^{ld} - y_b^{ld}\right\|^2\right\}. \tag{4}$$

where γ_b^{ld} is the inverse of the unknown noise variance at each NSCT decomposition level, l, and direction, d, of PAN image. Note that, with this modeling, we have decoupled each one of the bands of the contourlet transform and, since they are uncorrelated, we can do the estimation of each band independently of the other bands.

In the second stage of the hierarchical Bayesian framework we define the distribution on the parameters by using a gamma distribution

$$p(w|a_w, c_w) = \Gamma(w|a_w, c_w), \tag{5}$$

where $w > 0, w \in \Omega_b^{ld} = (\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld})$ denotes a hyperparameter, and $a_w > 0$ and $c_w > 0$ are, respectively, the shape and the inverse scale parameters of the distribution.

Finally, combining the first and second stages of the problem modeling, and defining $\Omega_b^{ld} = \{\alpha_b^{ld}, \beta_b^{ld}, \gamma_b^{ld}\}$, we have the global distribution

$$\begin{aligned} p(\Omega_b^{ld}, y_b^{ld}, x^{ld}, s_b^{ld}) =& p(\alpha_b^{ld})\, p(\beta_b^{ld})\, p(\gamma_b^{ld})\, p(y_b^{ld}|\alpha_b^{ld}) \\ & \times p(s_b^{ld}|y_b^{ld}, \beta_b^{ld})\, p(x^{ld}|y_b^{ld}, \gamma_b^{ld}), \end{aligned} \tag{6}$$

where $p(y_b^{ld}|\alpha_b^{ld})$, $p(s_b^{ld}|y_b^{ld}, \beta_b^{ld})$ and $p(x^{ld}|y_b^{ld}, \gamma_b^{ld})$ are given in Eqs. (2), (3), and (4), respectively.

The Bayesian paradigm dictates that inference on the parameters and the image, $(\Omega_b^{ld}, y_b^{ld})$, should be based on $p(\Omega_b^{ld}, y_b^{ld}|s_b^{ld}, x^{ld}) = p(\Omega_b^{ld}, y_b^{ld}, s_b^{ld}, x^{ld})/p(s_b^{ld}, x^{ld})$. Since $p(s_b^{ld}, x^{ld})$ cannot be calculated analytically, then $p(\Omega_b^{ld}, y_b^{ld}|s_b^{ld}, x^{ld})$ can not be found in closed form. We apply the variational methodology to approximate the posterior distribution by another distribution, $q(\Omega_b^{ld}, y_b^{ld})$, that minimizes the Kullback-Leibler(KL) divergence. We choose to approximate the posterior distribution $p(\Omega_b^{ld}, y_b^{ld}|s_b^{ld}, x^{ld})$ by the distribution $q(\Omega_b^{ld}, y_b^{ld}) = q(\Omega_b^{ld})q(y_b^{ld})$, where $q(y_b^{ld})$ and $q(\Omega_b^{ld})$ denote distributions on y_b^{ld} and Ω_b^{ld}, respectively.

The estimation of the parameters and the image is done iteratively. First, an estimation of each parameter $w \in \Omega_b^{ld}$ is selected as the mean of the posterior gamma distribution $q(w)$ and then the estimation of the Gaussian distribution of the HR MS coefficients, $q(y_b^{ld})$, is performed.

3. Classification

Once the pansharpened image has been obtained, its classification is carried out. The approach we follow (which will be described later) to improve the classification rate of one class will be tested on two classification methods which, for completeness, are briefly described now: linear discriminant analysis (LDA) and support vector machines (SVM).

LDA is an effective subspace technique that optimizes Fisher's score.[6] Subspace methods are a particular class of algorithms focused on finding projections of the original hyperdimensional space to a lower dimensional space where class separation is maximized. In addition, LDA does not require the tuning of free parameters. These good attributes have resulted in its extensive use and practical exploitation in remote sensing applications mainly focused on image classification and band selection. LDA is related to Fisher's linear discriminant and, roughly speaking, both aim at finding a linear combination of features that characterize or separate two or more classes.

SVM is one of the most successful examples of kernel methods, being a linear classifier that implements maximum margin separation between classes in a high dimensional Hilbert space $\mathcal{H}$. Kernel methods embed the data observed in the input space $\mathcal{X}$ into a higher dimensional space, the feature space $\mathcal{H}$, where the data are more likely to be linearly separable. Therefore, it is possible to build an efficient linear classifier in $\mathcal{H}$, that translates into a nonlinear classifier in the input space. The mapping function to perform such an embedding is denoted as $\mathbf{\Phi} : \mathcal{X} \rightarrow \mathcal{H}$. Computing the explicit mappings $\mathbf{\Phi}(\mathbf{x})$ of all the observed data points can be computa-

tional demanding, especially if the dimensionality of $\mathcal{H}$ is high. To avoid this problem and build efficient algorithms, kernel methods compute the similarity between training samples $\{\mathbf{x}_i\}_{i=1}^n$ using inner products between mapped samples instead of computing the dot product in the higher dimensional space explicitly. The so-called kernel matrix $\mathbf{K}_{ij} = K(\mathbf{x}_i, \mathbf{x}_j) = \langle \mathbf{\Phi}(\mathbf{x}_i), \mathbf{\Phi}(\mathbf{x}_j) \rangle$ contains all necessary information to perform many classical linear algorithms in the feature space, which are non-linear in the input space.[7]

It is important to note that, both for training and using the SVM for testing, one only needs to work with a valid kernel function, which should accurately reflect the similarity between samples. Valid kernels are functions representing a dot product in $\mathcal{H}$. The radial basis function (RBF), $K(\mathbf{x}, \mathbf{z}) = \exp\left(-\|\mathbf{x} - \mathbf{z}\|^2 / 2\sigma^2\right)$, $\sigma \in \mathbb{R}^+$ was the kernel function selected in this work. To implement SVM for multiclass problems we used the one-versus-all strategy given the particular characteristics of the proposed scheme.

4. Improving the Classification Performance for a Single Class

Once the pansharpening method described in section 2 has been used on a LR MS image and one of the classification methods described above has been applied, the user may be interested in boosting the performance of the classifier on a given class. In this section we propose to recalculate the parameters of the pansharpening method in order to obtain a new pansharpened image with an improved classification rate for the class of interest.

By examining the HR classified image, both visually and numerically (using for instance the confusion matrix), the user selects a class to improve its classification figures of merits. A new estimation of the image and parameters is performed. Utilizing the already estimated pansharpened image, the parameters for the new reconstruction are estimated utilizing only the pixels belonging to the class of interest in this image. Using those parameters a new pansharpened image is obtained. No iteration between parameter and image estimates is required.

This result in an estimation of the image whose spectral and spatial characteristics are more tailored to the pixels in the class of interest and, hence, will hopefully increase the classification performance for the elements of the class. Note however that this may imply, as we will see in the experimental section, that the classification performance on the other classes

may decrease.

5. Experimental Results

Experiments were run on a Quickbird image. The MS image, depicted in real color in Figure 1a, has a spatial resolution of 256×256 pixels with each pixel covering a square area with a side of 2.4 m and four spectral bands: blue (450-520 nm), green (520-600 nm), red (630-690 nm), near-IR (760-900 nm). The PAN image (see Fig. 1b) has a resolution of 1024×1024 pixels with a size of 0.6 m covering the whole spectral interval (405-1053 nm). The result of the pansharpening process, with the parameters automatically estimated using all the MS and PAN images, is shown in Fig. 1d.

Using the MS and PAN images, a small number of pixels were classified into ten different classes (cars, water, forest, ...). This set of pixels, depicted in Fig. 1c, is considered our ground truth. We randomly chose 20% of the samples of each class to train the LDA and SVM classifiers and the rest was used for testing. In order to incorporate both spectral and spatial characteristics for each pixel into the classification process, we used a descriptor composed of the value of each pixel under consideration and its four nearest neighbors. Since each pixel has associated five values, four corresponding to the MS bands and another one for the panchromatic, the descriptor for each pixel has 25 components.

The classification quality is measured using the precision and recall values on a given class defined as

$$recall = \frac{TP}{TP + FN}, \quad precission = \frac{TP}{TP + FP} \tag{7}$$

where TP is the number of pixels in class correctly classified, FN is the number of pixels in the class incorrectly classified and FP is number of pixels not belonging to a class incorrectly classified as belonging to the class. Table 1 presents the figures of merit for each classifier on the pansharpened image in Fig. 1d. This image presents a very high level of spatial detail with no chromatic distortion. The classification figures show that SVM outperforms LDA although both classifiers perform well for all the classes except classes 10 and 8 where they perform poorly, especially the LDA classifier.

A class is now selected to improve its classification figures. In this case class 10 (isolated tree) was selected although similar results were obtained when selecting the other classes. Using only the pixels of the MS and PAN image belonging to the selected class, the parameters were estimated us-

ing the procedure described in section 4 and a new pansharpened image, depicted in Fig. 1e, was obtained.

Using this image, the classifiers were trained and a new classification step was performed obtaining the results presented in Table 2. Although the reconstructed images using the parameters estimated from the whole image (Fig. 1d) and the parameters estimated using only the pixels of the class 10 (Fig. 1e) are very similar from a visual point of view, the classification figures show a higher precision and recall for the selected class 10 and, also, for many others. Note however, that some classes, like classes 6 or 8, perform slightly worse with those parameters.

Table 1. Recall and Precision values obtained using the pansharpened image with parameters estimated from all the pixels in the image.

Class	LDA recall	LDA precision	SVM recall	SVM precision
1. Asphalt	0.89	0.91	0.99	0.98
2. Dense Forest	0.75	0.72	0.91	0.93
3. Forest	0.87	0.98	0.99	0.98
4. Bare Soil	0.93	0.86	0.99	0.99
5. Building	0.84	0.93	0.99	0.98
6. Grass	0.82	0.82	0.96	0.94
7. Dry Grass	0.99	0.77	0.99	0.99
8. Car	0.63	0.66	0.81	0.98
9. Water	0.93	0.74	0.97	0.98
10. Isolated Tree	0.89	0.29	0.82	0.89

6. Conclusions

In this paper we have shown that pansharpening techniques can be used to increase the performance of classification methods when are applied to MS images. We have addressed the problem of adaptively modifying a pansharpening method in order to improve the precision and recall figures of merit of the classification on a given class without deteriorating the performance of the classifier over the other classes. The validity of the proposed technique has been demonstrated using a real Quickbird image. Work is being currently carried out to theoretically justify the used approach.

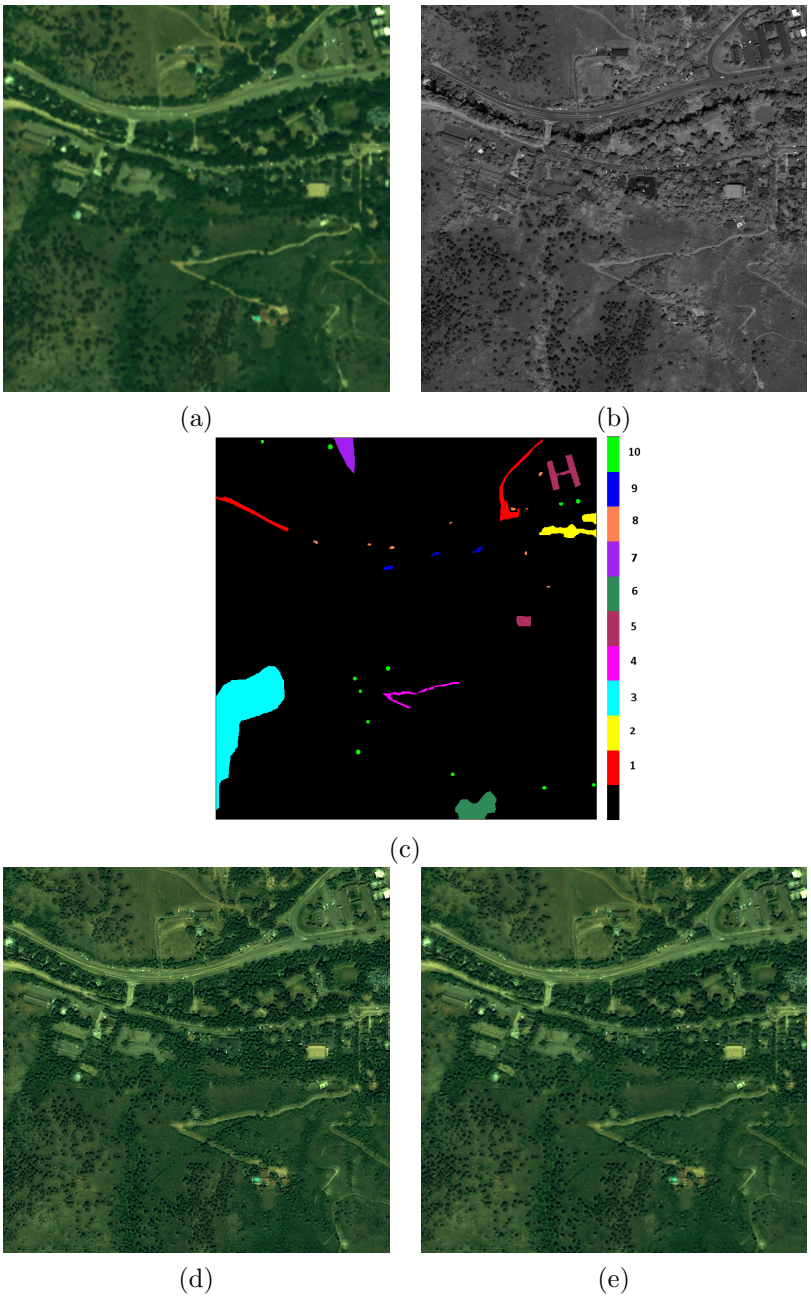

Fig. 1. (a) MS and (b) PAN images. (c) Ground truth. (d) Pansharpened image using the super resolution method described in section 2 with the parameters estimated using the whole image. (e) Pansharpened image utilizing only the pixels of the training set in class 10 to estimate the model parameters.

Table 2. Recall and Precision values obtained using the pansharpened image with parameters estimated only from pixels of the class 10.

Class	LDA recall	LDA precision	SVM recall	SVM precision
1. Asphalt	0.89	0.90	0.99	0.97
2. Dense Forest	0.72	0.73	0.92	0.93
3. Forest	0.88	0.99	0.99	0.99
4. Bare Soil	0.94	0.87	0.99	0.99
5. Building	0.82	0.93	0.99	0.98
6. Grass	0.83	0.80	0.96	0.95
7. Dry Grass	0.99	0.80	0.99	0.99
8. Car	0.54	0.56	0.75	0.94
9. Water	0.94	0.72	0.99	0.99
10. Isolated Tree	0.89	0.30	0.88	0.92

Acknowledgements

This work has been supported by the Consejería de Innovación, Ciencia y Empresa of the Junta de Andalucía under contract P07-FQM-02701, by the Comisión Nacional de Ciencia y Tecnología under contract TIN2010-15137, and the Spanish research program Consolider Ingenio 2010: MIPRCV (CSD2007-00018).

References

1. I. Amro, Multispectral Image fusion using Multiscale and Super-resolution methods, PhD thesis, Dept. of Computer Science and Artificial Intelligence, Universidad de Granada 2011.
2. L. Bruzzone, L. Carlin, L. Alparone, S. Baronti, A. Garzelli and F. Nencini, Can multiresolution fusion techniques improve classification accuracy? *Image and Signal Processing for Remote Sensing XII*, **6365** 2006.
3. I. Amro, J. Mateos, M. Vega, General contourlet pansharpening method using Bayesian inference, in *2010 European Signal Processing Conference (EUSIPCO-2010)*, 2010.
4. A. L. da Cunha, J. Zhou and, M. N. Do, The nonsubsampled contourlet transform: theory, design, and applications, in *IEEE Trans. Image Proc.*, **15**, 3089 (2006).
5. M. Vega, J. Mateos, R. Molina and A.K.. Katsaggelos, Super resolution of multispectral images using TV image models, in *2th Int. Conf. on Knowledge-Based and Intelligent Information & Eng. Sys.*, 2008.
6. R. Duda and P. Hart, *Pattern classification and scene analysis* (Wiley, New York, USA, 1973).
7. J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*, (Cambridge University Press, 2004).