FUSION OF THERMAL INFRARED HYPERSPECTRAL AND VIS RGB DATA USING GUIDED FILTER AND SUPERVISED FUSION GRAPH

Wenzhi Liao¹, Frieke Van Coillie², Sidharta Gautama¹, Aleksandra Pižurica¹, Wilfried Philips¹

¹ Ghent University-TELIN-IPI-iMinds, Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium
Email: {Wliao, Sanja, Sidharta.Gautama, Philips}@telin.ugent.be
² FORSIT Research Unit, Ghent University, Coupure Links 653, 9000 Ghent, Belgium
Email: Frieke.Vancoillie@UGent.be

ABSTRACT

Nowadays, advanced technology in remote sensing allows us to get multi-sensor and multi-resolution data from the same region. Fusion of these data sources for classification remains challenging problems. We proposed a novel framework for fusion of low spatial-resolution Thermal Infrared (TI) hyperspectral (HS) and high spatial-resolution RGB data. First, we do super-resolution on TI HS data by using RGB image and guided filter in principal component analysis (PCA) domain. Then, we couple feature extraction and data fusion of spectral features (from TI HS data) and spatial features (morphological features generated on RGB data) through supervised fusion graph. Finally, the fused features are used by SVM classifier to generate the final classification map. Experimental results on the classification of fusing real TI HS and RGB images demonstrate the effectiveness of the proposed method both visually and quantitatively.

Index Terms— Data fusion, multi-resolution, remote sensing, hyperspectral image

1. INTRODUCTION

Recent advances in the sensors technology of remote sensing (RS) have led to an increased availability of acquiring multi-sensor and multi-resolution data from the same area. In particular, hyperspectral (HS) images provide a detailed description of the spectral signatures of ground covers but with low spatial resolution, whereas visual RGB images with high resolution give detailed spatial information the same surveyed area. Many techniques have been developed for fusion of multi-sensor and multi-resolution RS imagery [1-5]. To super-resolve the low spatial resolution HS/multispectral (MS) to the same spatial size of high resolution RGB/panchromatic (PAN) image, some of these approaches employ the so-called component substitution methods [1] or their generalization [2]. Others model PAN image as a linear combination of the ideal MS bands, and restore an ideal high-resolution MS image by utilizing different regularization [3]. Recently, Licciardi et al. [4] combined non-linear PCA (NLPCA) and Indusion to enhance the spatial resolution of the HS image by fusing a PAN image. Their method first applied NLPCA to project the original HS data into a lower space, then enhanced the derived non-linear PCs by the Indusion process; finally, got the high spatial resolution HS data by inverting NLPCA. However, these approaches either suffer from spectral distortions or from high cost on computational time to estimate a good solution.

For classification tasks, Thoonen et al. [5] proposed composite decision fusion to fuse the classification maps obtained from both a low spatial HS image and a high spatial RGB image using color attribute profiles. However, their method super-resolved HS by just using simply cubic interpolation, which will cause spatial distortions in the classification map. Our graph-based data fusion method won the “Best Paper Challenge” award of 2013 IEEE Data Fusion Contest, but with unsupervised fusion graph [6], which does not take the class discrimination into account.

In this paper, we propose a novel framework for both super-resolution and data fusion of a low spatial resolution Thermal Infrared (TI) HS image and a high spatial RGB image of the same scene. Fig. 1 shows our proposed framework. We first use RGB image and guided filter [7] to do super-resolution of TI HS image in PCA domain, and generate morphological features on original RGB image. Then, we couple dimensionality reduction and data fusion of the spectral information (of super-resolved TI HS image) and the morpholog-
2. MORPHOLOGICAL FEATURES OF RGB IMAGE

Morphological features are generated by either applying morphological openings or closings by reconstruction on the image, using a structural element (SE) of predefined size. An opening acts on bright objects compared with their surrounding, while closings act on dark objects. For example, an opening deletes bright objects that are smaller than the SE. By increasing the size of the SE and repeating the previous operation, a complete morphological profile (MP) is built, carrying information about the size and the shape of objects in image.

In our experiments, morphological features are generated by applying morphological openings and closings with partial reconstruction on RGB image. The effect of using morphological features with partial reconstruction for classification of remote sensing data from urban areas has been discussed in our previous work [6,9]. MPs with 10 openings and closings (ranging from 1 to 10 with step size increment of 1) are computed on RGB image with disk SE.

3. PROPOSED METHOD FOR SUPER-RESOLUTION AND DATA FUSION

3.1. Proposed super-resolution by PCA and guided filter

One of the main challenges of fusion a low spatial HS and a high resolution RGB to get a high spatial resolution HS, is not easy to make a balance on spectral and spatial preservation. Recently, the guided filter [7] has been widely used in many applications (e.g. edge-aware smoothing, detail enhancement and etc.), as its efficient and strong abilities to transfer the structures of the guidance image to the filtering output. Its application to HS data can be found in [8].

In this paper, we propose a novel method to enhance the spatial resolution of TI HS image by using PCA and guided filter. Instead of component substitution which may cause spectral distortions, we use a high resolution RGB image to guided filter the super-resolution of low spatial resolution HS image. By guided filtering the super-resolution process, our method can not only preserve the spectral information from the original HS image, but also transfer the spatial structures of high resolution RGB image to the enhanced HS image. To speed up the processing time, our method first uses PCA to decorrelate the HS images and separate the information content from the noise. The first $k$ PCA channels contain most of
problems like the curse of dimensionality, excessive computation time and so on. We proposed a graph-based data fusion method [9] to fuse hyperspectral and LiDAR data for classification, but with unsupervised fusion graph. In this paper, we extend it to supervised version, which take into account class discrimination. Let $X^{HS} = \{x^{HS}_i\}_{i=1}^n$, $X^{MPS} = \{x^{MPS}_i\}_{i=1}^n$ and $Y = \{y_i\}_{i=1}^n$ denote the spectral features after our proposed super-resolution, MPs computed on original RGB image, and the class labels (where $y_i \in \{1, \ldots, C\}$, $C$ is the number of classes), respectively. $X^{Sta} = \{x^{Sta}_{i}\}_{i=1}^n = [X^{HS};X^{MPS}]$ denotes the vector stacked by the spectral and spatial features.

We assume all data sources are scaled to the same ranges before fusion. We first build a graph for each data source, for example, the graph constructed by our super-resolved HS (i.e., $G^{HS} = (X^{HS}, A^{HS})$), where $A^{HS}$ represents the edges of the graph. The edge between data point $x^{HS}_i$ and $x^{HS}_j$ is here denoted as $A^{HS}_{ij} \in \{0, 1\}$; $A^{HS}_{ij} = 1$ if $x^{HS}_i$ and $x^{HS}_j$ are “close” and $y_i = y_j$, whereas $A^{HS}_{ij} = 0$ if $x^{HS}_i$ and $x^{HS}_j$ are “far apart” or $y_i \neq y_j$. The “close” is defined as belonging to $K$ nearest neighbors ($KNN$) of the other data points. The $KNN$s of the data point $x^{HS}_i$ are its $K$ nearest neighbors in terms of spectral signatures. On the other hand, when the graph is constructed by MPs, the $KNN$s of the data point $x^{MPS}_i$ are its $K$ nearest neighbors in terms of spatial characteristics. We define a fusion graph $G^{Fus} = (X^{Sta}, A^{Fus})$, where $A^{Fus} = A^{HS} \odot A^{MPS}$. The operator ‘$\odot$’ denotes element-wise multiplication, i.e. $A^{Fus}_{ij} = A^{HS}_{ij} A^{MPS}_{ij}$.

This means that the stacked data point $x^{Sta}_i$ and $x^{Sta}_j$ are connected only if they have similar both spectral and spatial characteristics and they are from the same class. For more details to obtain the fused features, we refer the readers to our recent work reported in [6].

### 4. EXPERIMENTAL RESULTS

Experiments are done on a thermal infrared hyperspectral data and a visual RGB image which were acquired by Telops Inc. on May 2013 over an urban area near Thetford Mines in Qu´ebec, Canada. The TI HS image has 84 spectral bands that covers the wavelengths between 7.8 to 11.5 µm with approximately 1-m spatial resolution. The visible RGB image is a series of color images acquired during separate flight-lines with approximately 20-cm spatial resolution. The whole scene of both data contains 7 classes, but with different spatial size of which the TI HS consists of 874×751 pixels while RGB of 4386×3769.

The SVM classifier with radial basis function (RBF) kernels is applied in our experiments. We apply a grid-search on the two parameters $C$ and $\gamma$ of SVM using 5-fold cross-validation to find the best $C$ within the given set $\{10^{-1}, 10^{0}, 10^{1}, 10^{2}, 10^{3}\}$ and the best $\gamma$ within the given set $\{10^{-3}, 10^{-2}, 10^{-1}, 10^{0}, 10^{1}\}$. We compare our proposed fusion method (SGF) with the schemes of (1) Simply enlarging the original HS image by cubic interpolation (Cub); (2) PCA component substitution method (PCA), similar as [1]; (3) Our proposed image fusion using guided filter in PCA domain (GFP); (4) Using MPs computed on original RGB image (MPs); (5) Stacking our enhanced HS image and MPs (Sta); (6) Unsupervised graph-based fusion method (UGF) [6] to fuse our enhanced HS image and MPs. For quantitative comparisons, we divide all labeled samples spatially isolated into two equal groups, one group for training set, the other for test. Within the training set, we randomly select 1000 samples per class for training, the results are averaged over five runs. The classification results are quantitatively evaluated by measuring the Overall Accuracy (OA), the Average Accuracy (AA) and the Kappa coefficient ($\kappa$) on the test data. The experiments were carried out on 64-b. 3.40 GHz Intel i7-4930K (1 core) CPU computer with 64 GB memory, the consumed time includes image fusion, feature fusion and classification. Table 1 shows the accuracies and consumed time (hours) obtained from the experiments, Fig. 3 shows the best result of each method.

<table>
<thead>
<tr>
<th>Fcat.</th>
<th>Cub</th>
<th>PCA</th>
<th>GFP</th>
<th>MPs</th>
<th>Sta</th>
<th>UGF</th>
<th>SGF</th>
</tr>
</thead>
<tbody>
<tr>
<td>OA (%)</td>
<td>78.1</td>
<td>78.4</td>
<td>91.6</td>
<td>88.7</td>
<td>90.6</td>
<td>93.2</td>
<td>96.7</td>
</tr>
<tr>
<td>AA (%)</td>
<td>72.4</td>
<td>73.9</td>
<td>84.7</td>
<td>86.4</td>
<td>87.3</td>
<td>90.7</td>
<td>96.6</td>
</tr>
<tr>
<td>$\kappa$ (%)</td>
<td>72.8</td>
<td>73.8</td>
<td>89.5</td>
<td>87.8</td>
<td>88.9</td>
<td>91.5</td>
<td>95.9</td>
</tr>
<tr>
<td>Time (h)</td>
<td>3.81</td>
<td>2.82</td>
<td>1.17</td>
<td>0.67</td>
<td>6.86</td>
<td>0.61</td>
<td>0.53</td>
</tr>
</tbody>
</table>
classification.

5. CONCLUSION

The contribution of this paper is a methodology to fuse a low spatial resolution HS image and a high resolution RGB image in the classification task. Some existing image fusion methods suffer either spectral distortions or spatial distortions. Our proposed image fusion method can preserve both spectral and spatial information by using RGB image and guided filter in PCA domain. Our supervised graph-based fusion considers the class discrimination to couple dimension reduction and data fusion, it can make full advantage of each data sources and reduce the computational cost for classification. Experimental results on the classification of the real TI HS and RGB images show the efficiency of the proposed method.

Acknowledgment

This work was financially supported by the SBO-IWT project Chameleon: Domain-specific Hyperspectral Imaging Systems for Relevant Industrial Applications. The authors would like to thank Telops Inc. (Québec, Canada) for acquiring and providing the data used in this study, the IEEE GRSS Image Analysis and Data Fusion Technical Committee and Dr. Michal Shimony (Signal and Image Centre, Royal Military Academy, Belgium) for organizing the 2014 Data Fusion Contest, the Centre de Recherche Public Gabriel Lippmann (CRPGL, Luxembourg) and Dr. Martin Schlerf (CRPGL) for their contribution of the Hyper-Cam LWIR sensor, and Dr. Michaela De Martino (University of Genoa, Italy) for her contribution to data preparation.

6. REFERENCES