A TOPOLOGICAL DATA ANALYSIS GUIDED FUSION ALGORITHM: MAPPER-REGULARIZED MANIFOLD ALIGNMENT

Jingliang Hu[1,2], Danfeng Hong[1,2], Yuanyuan Wang[2], and Xiao Xiang Zhu[1,2]

[1] *Remote Sensing Technology Institute (IMF), German Aerospace Center (DLR), Germany*
[2]*Signal Processing in Earth Observation (SiPEO), Technische Universität München (TUM), Germany*

ABSTRACT

Hyperspectral images and polarimetric synthetic aperture radar (PolSAR) data are two important data sources, yet they barely appear under the same scope, even though multi-modal data fusion is attracting more and more attention. To our best knowledge, this paper investigates for the first time semi-supervised manifold alignment (SSMA) for the fusion of the hyperspectral image and PolSAR data. The SSMA searches a latent space where different data sources are aligned, which is accomplished by using the label information and the topological structure of the data. This paper is the first attempt to apply topological data analysis (TDA), a recent mathematic sub-field of data analysis, in remote sensing. It aims to reveal relevant information from the shape of a data in its feature space, and has been proven powerful in medicine. The paper also proposes a novel algorithm, MAPPER-regularized manifold alignment, which embeds the TDA into a semi-supervised manifold alignment for the fusion of the hyperspectral image and PolSAR data. The proposed algorithm exhibits superior performance in fusing a simulated EnMAP data set and a Sentinel-1 data set for an image of Berlin.

Index Terms— Classification, data fusion, EnMAP, hyperspectral image, land cover, land use, manifold alignment, MAPPER, PolSAR, semi-supervised learning, Sentinel-1, topological data analysis (TDA).

1. INTRODUCTION

Multimodal data fusion in remote sensing has recently attracted a lot of attention, because the complementary information derived from different data sources promotes the performance of applications [1–3]. The high spectral resolution of the hyperspectral image is an ideal resource for remote sensing data analysis [4–6]. Polarimetric synthetic aperture radar (PolSAR) data provides the dielectic property, orientation, and geometric shape of an observed material. Therefore, it could be beneficial to fuse the hyperspectral image and PolSAR data [7]. The manifold alignment technique has successfully accomplished data fusion in both the machine learning field and remote sensing applications [8–10]. In this work, we investigate the manifold alignment approach for the fusion of the hyperspectral image and PolSAR data. This technique leverages the label information and the topological structure of the data to find a latent space where multimodal data can be aligned to each other. Among manifold techniques [8, 10], the k-nearest-neighbor is commonly used for the approximation of the topological structure of a data. However, recently an emerging mathematical sub-field, topological data analysis (TDA), has been developed from the applied topology. It aims to retrieve relevant information from the topological structure of a data set. It has already been proven in medicine that one of the TDA tools, MAPPER [11], is capable of revealing valuable unknown knowledge [12]. MAPPER offers two advantages. First, the MAPPER is able to incorporate field knowledge while deriving the topological structure of a data. Second, the MAPPER derives a regional-to-global topological structure that is robust to outliers in a feature space.

Inspired by studies of medicine and a comprehensive theoretical foundation [11], we propose a novel MAPPER-regularized manifold alignment algorithm for the fusion of the hyperspectral image and PolSAR data. The contributions of this study are three-fold:

- To the best of our knowledge, this study is the first attempt to fuse heterogeneous remote sensing data sources, namely, hyperspectral data and PolSAR data, using the manifold alignment technique.

- It is also the first time to explore the topological data analysis (TDA) technique in remote sensing.

- A novel MAPPER-regularized manifold alignment algorithm is proposed for the fusion of the hyperspectral image and the PolSAR data.

2. METHODOLOGY

2.1. MAPPER

In order to introduce MAPPER in an understandable way, we illustrate the MAPPER with an intuitive example, as shown in Fig. 1. For more mathematical details, please refer to [11].

The MAPPER is a tool developed from applied topology for analyzing and visualizing big data sets. It essentially re-
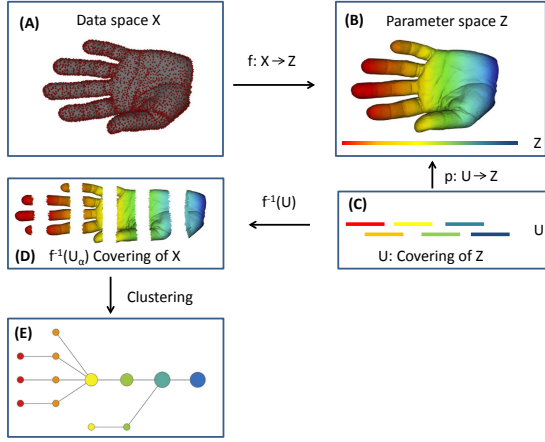
Fig. 1: The workflow of the MAPPER with a point cloud of a hand. (A): Data space X, point cloud of a hand; (B): Parameter space $Z = fX$; Filter function f: distance to the wrist; (C): the overlapped bins; (D): $f^{-1}(U_\alpha)$ covering of X; (E) Topological structure, achieved by clustering slices of data. Modified from [13].

trieves relevant information about the shape that a data forms in feature space. The MAPPER consists of three components:

- Filter function. An application-oriented filter function needs to be chosen for the MAPPER, as the first step. This selected filter defines a perspective where an expert would interpret the data. As shown in Fig. 1 (B), the filter is defined as the distance to the wrist.

- Data separation. After filtering, the continuous value range is sliced into overlapped bins [14] with a given overlap percentage and a number of bins, shown in Fig. 1 (C). The original input data can thus be sliced into data bins accordingly, as shown in Fig. 1 (D).

- Clustering and visualization construction. Clustering is applied on each of the sliced data bins. Clusters of adjacent data bins might share the same data points. MAPPER represents the topological structure with a graph where a node represents a cluster, and an edge represents a link of two clusters. A link is generated for two clusters if they share same data points. The graph serves as a simplified visualization of the topological structure of a data set, as shown in Fig. 1.

2.2. MAPPER-regularized Manifold Alignment

Let $\mathbf{X}_i = [\mathbf{x}_i^1, ..., \mathbf{x}_i^k, ..., \mathbf{x}_i^{n_i}] \in \mathbb{R}^{m_i \times n_i}$ be a matrix representing the ith data source, with m_i dimensions by n_i instances. The term $\mathbf{x}_i^k$ denotes the kth instance of the ith data source. Let K denote the total number of data sources. A semi-supervised manifold alignment (SSMA) aims to learn a set of K projections $\{f_1, ..., f_K\}$ so that the ith projection f_i

maps the ith data source $\mathbf{X}_i$ into the latent space, where the K data sources are aligned in terms of three desired properties: **(a)**, the data of same class are located close to each other, which is mathematically presented as a similarity matrix (1);

$$W_s = \begin{pmatrix} W_s^{1,1} & ... & W_s^{1,K} \\ ... & ... & ... \\ W_s^{K,1} & ... & W_s^{K,K} \end{pmatrix} \qquad (1)$$

(b): data of different classes are located far from one another, formulated as a dissimilarity matrix (2);

$$W_d = \begin{pmatrix} W_d^{1,1} & ... & W_d^{1,K} \\ ... & ... & ... \\ W_d^{K,1} & ... & W_d^{K,K} \end{pmatrix} \qquad (2)$$

(c): the topological property of individual data is preserved, presented as a topology matrix (3).

$$W_t = \begin{pmatrix} W_t^{1,1} & 0 & 0 \\ 0 & ... & 0 \\ 0 & 0 & W_t^{K,K} \end{pmatrix} \qquad (3)$$

Each of the matrices (1), (2), and (3) is a matrix with the size of $(n_1 + n_2 + ... + n_k) \times (n_1 + n_2 + ... + n_k)$. In each matrix, $W^{i,j}$ is a matrix representing the relationship between the ith and j^{th} data sources on the individual property. If x_i^p and x_j^q share the same label, then $W_s^{i,j}(p,q) = 1$; otherwise $W_s^{i,j}(p,q) = 0$. If x_i^p and x_j^q belong to different classes, then $W_d^{i,j}(p,q) = 1$; otherwise $W_s^{i,j}(p,q) = 0$. $W_t^{i,i}(p,q) = 1$ if x_i^p and x_i^q are neighbors. Otherwise, $W_t^{i,i}(p,q) = 0$.

Three terms $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{C}$, are formulated in terms of properties **(a)**, **(b)**, and **(c)**, respectively, so that the cost function $\mathcal{L}$ could be constructed for an optimazition problem.

$$\mathcal{A} = \sum_{i=1}^{K} \sum_{j=1}^{K} \sum_{p=1}^{n_i} \sum_{q=1}^{n_j} \|f_i^T x_i^p - f_j^T x_j^q\|^2 W_s^{i,j}(p,q). \qquad (4)$$

$$\mathcal{B} = \sum_{i=1}^{K} \sum_{j=1}^{K} \sum_{p=1}^{n_i} \sum_{q=1}^{n_j} \|f_i^T x_i^p - f_j^T x_j^q\|^2 W_d^{i,j}(p,q). \qquad (5)$$

$$\mathcal{C} = \sum_{i=1}^{K} \sum_{p=1}^{n_i} \sum_{q=1}^{n_i} \|f_i^T x_i^p - f_i^T x_i^q\|^2 W_t^{i,i}(p,q). \qquad (6)$$

$$\mathcal{L}(f_1, ..., f_K) = (\mathcal{A} + \mathcal{C})/\mathcal{B} \qquad (7)$$

$$\underset{f_1,...,f_K}{\operatorname{argmin}} \mathcal{L}(f_1, ..., f_K) \qquad (8)$$

The solution $f_1, ..., f_K$ that minimizes the cost function $\mathcal{L}(f_1, ..., f_K)$ is given by the smallest non-zero eigenvector of the generalized eigenvalue decomposition of Eq. (9). And the matrix D and the matrix L in Eq. (9) are the degree matrix and the Laplacian matrix, respectively.

$$Z(\mu L_t + L_s)Z^T x = \lambda Z L_d Z^T x, \qquad (9)$$

where

$$Z = \begin{pmatrix} X_1 & 0 & ... & 0 \\ ... & ... & ... & ... \\ 0 & ... & 0 & X_K \end{pmatrix},$$

$$L_a = W_a - D_a, a \in \{s, d, t\}$$

$$D_a(p, q) = \begin{cases} \sum_{q=1}^{m_1 + ... + m_k} W_a(p, q) & p = q \\ 0 & p \neq q \end{cases}.$$

Rather than approximating the topological structure by *kNN*, we use MAPPER to extract the topological structure in the proposed algorithm. The topological structure of data source $\mathbf{X}_i$ could be represented as an $n_i \times n_i$ matrix W_c^i, where n_i is the number of instances: $W_c^i(p, q) = 1$, when data instances p and q are in the same cluster or in linked clusters; otherwise, $W_c^i(p, q) = 0$. Thus, the topological matrix W_t in equation (3) is replaced by W_c (10).

$$W_c = \begin{pmatrix} W_c^{1,1} & 0 & 0 \\ 0 & ... & 0 \\ 0 & 0 & W_c^{K,K} \end{pmatrix}. \quad (10)$$

3. EXPERIMENT AND DISCUSSION

To demonstrate the performance of the proposed algorithm, a region in western Berlin is selected for the experiment. Fig. 2 demonstrates the simulated EnMAP data, the Sentinel-1 dual-Pol SAR data, the training data, and the testing data of the selected region. The simulated EnMAP data includes a scene with a size of 817 by 220, a 30-meter ground sampling distance (GSD), and 244 spectral bands ranging from $400\,\mathrm{nm}$ to $2500\,\mathrm{nm}$ [15]. The Sentinel-1 data is a VH-VV polarized dual-Pol SAR data set with a GSD of 13 meters and a size of 1723 by 476. The training data and the testing data are a land cover land use data set derived from the Open Street Map[1].

To design the input like [16], the first four principal components (PCs) are extracted from the hyperspectral image, which accounts for 99% of the variances. The morphological profile is then extracted from these four PCs, with radius equal to one, two, and three. In total, 28 features are extracted from the simulated EnMAP data set. For the Sentinel-1 data, four polarimetric features are derived: intensity of the VH channel, intensity of the VV channel, the coherence of VV and VH, and the intensity ratio of VV and VH. The morphological profile was also used to extract spatial information from dual-Pol data with radius equal to one, two, and three. In total, it results in 28 features from Sentinel-1 dual-Pol data.

The proposed algorithm is evaluated by comparing it to two other data fusion algorithms and individual usage of the two data sources, in terms of classification. Thus, five groups of classification experiments have been done, which are: (*A*)
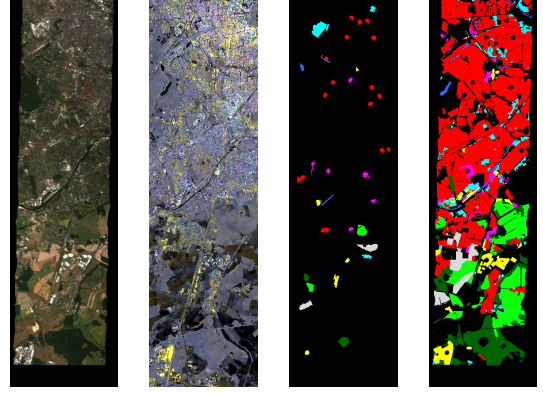
Fig. 2: Experiment data for the city of Berlin, Germany. From left to right: Simulated EnMAP data; Sentinel-1 dual-Pol data; Training set; Testing set.

dual-Pol SAR data (POL), (*B*) optical data (OPT), (*C*) fusing of optical and dual-Pol SAR data by feature concatenation (OPT-POL), (*D*) fusing optical and dual-Pol SAR data by SSMA [8], and (*E*) fusing optical and dual-Pol SAR data by the proposed algorithm (TOPMA).

For parameter tuning, the *k* of the *kNN* is set as nine in SSMA, according to the recommendation in [8]. The dimension of the latent space and the μ in Eq. 8 will be discussed in experiments. In TOPMA, the number of intervals for dividing the data is set as five, and the overlapping rate is set as 50%. The first and second principal components are chosen to serve as the filter function for the TOPMA.

For classification, three classical classifiers are chosen: the one-nearest-neighbor classifier (*ONE-NN*), the linear support vector machine (*LSVM*), and the kernel support vector machine (*KSVM*). In this work, hyperparameter tuning of *LSVM* and *KSVM* automatically take place with MATLAB function as used in [17].

Fusion vs. non-fusion. According to the experiment results in Fig. 3, data fusion improves the performance of classification when compared to classification on individual data source. It proves that fusion of the EnMAP data and the Sentinel-1 data is advantageous to classification.

MA-fusion vs. non-MA-fusion. MA-fusion refers to manifold alignment fusion, the SSMA and the TOPMA. Non-MA-fusion includes the OPT-POL, a feature concatenation fusion. According to Fig. 3, among all three classifiers, the proposed TOPMA has the best performance. And it is obvious that the MA-fusion outperforms the OPT-POL. An exception is that the OPT-POL outperforms the SSMA when *KSVM* is applied in classification. The *KSVM* decides classification hyperplane boundaries in a space where the data was non-linearly projected with the Gaussian kernel. For the projection, a kernel-based data fusion process is carried out. The kernel-based data fusion introduces uncertainties into the experiments for analysis. However, in the same scenario, the
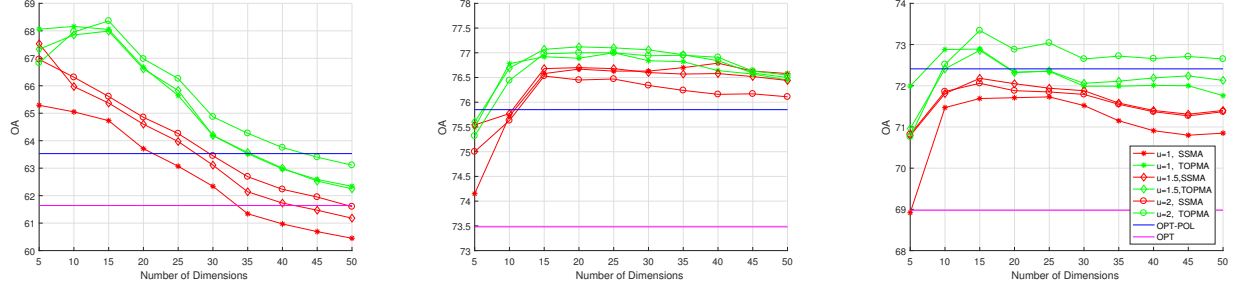
Fig. 3: The charts show the classification performance in terms of overall accuracy (OA) for the experiments applied on data from Berlin, Germany. From left to right, they show results from *ONE-NN*, *LSVM*, and *KSVM*, respectively. The OA is reported as percentage. The OA of POL are not shown in the charts; they are 43% (*ONE-NN*), 58% (*LSVM*), and 52% (*KSVM*).

TOPMA still provides superior performance.

TOPMA vs. SSMA. It is obvious in Fig. 3 that, when parameter μ and the dimension of latent space are the same for TOPMA and SSMA, TOPMA outperforms SSMA.

Parameter μ. The parameter μ is a balancing weight between the topological structure and the label information while using MA-based data fusion. A larger value of μ assigns greater weight to topological structure. According to Fig. 3, a larger value of μ in SSMA and TOPMA results in better classification performance than a smaller one.

4. CONCLUSION

We propose a novel algorithm for the fusion of the hyperspectral image and polarimetric SAR data. It couples a topological data analysis tool with the manifold alignment technique. The proposed fusion algorithm exhibits superior performance in an experiment of fusing a scene of the simulated EnMAP data and the Sentinel-1 dual-pol SAR data.

Future work will compare the proposed algorithm with more fusion algorithms and test on more data sets. We will also investigate the impact of the filter function, the number of bins, and the overlap rate in this novel fusion algorithm.

5. REFERENCES

[1] M. Dalla Mura, S. Prasad, F. Pacifici, P. Gamba, J. Chanussot, and J. A. Benediktsson, "Challenges and opportunities of multimodality and data fusion in remote sensing," *Proc. IEEE*, vol. 103, no. 9, pp. 1585–1601, 2015.

[2] J. Hu, L. Mou, A. Schmitt, and X. Zhu, "Fusionet: A two-stream convolutional neural network for urban scene classification using polsar and hyperspectral data," in *JURSE, 2017*.

[3] X. Liu, C. Deng, J. Chanussot, D. Hong, and B. Zhao, "StfNet: A two-stream convolutional neural network for spatiotemporal image fusion," *IEEE Trans. Geosci. Remote Sens.*, 2019.

[4] D. Hong, N. Yokoya, J. Xu, and X. Zhu, "Joint & progressive learning from high-dimensional data for multi-label classification," in *Proc. ECCV*, 2018, pp. 469–484.

[5] D. Hong, N. Yokoya, J. Chanussot, and X. Zhu, "An augmented linear mixing model to address spectral variability for hyperspectral unmixing," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1923–1938, 2019.

[6] R. Hang, Q. Liu, D. Hong, and P. Ghamisi, "Cascaded recurrent neural networks for hyperspectral image classification," 2019.

[7] J. Hu, P. Ghamisi, and X. Zhu, "Feature extraction and selection of sentinel-1 dual-pol data for global-scale local climate zone classification," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 9, pp. 379, 2018.

[8] D. Tuia, M. Volpi, M. Trolliet, and G. Camps-Valls, "Semisupervised manifold alignment of multimodal remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 12, pp. 7708–7720, 2014.

[9] D. Hong, N. Yokoya, J. Chanussot, and X. Zhu, "Cospace: Common subspace learning from hyperspectral-multispectral correspondences," *IEEE Trans. Geosci. Remote Sens.*, 2019.

[10] D. Hong, N. Yokoya, N. Ge, J. Chanussot, and X. Zhu, "Learnable manifold alignment (LeMA): A semi-supervised cross-modality learning framework for land cover and land use classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 147, pp. 193–205, 2019.

[11] Gurjeet Singh, Facundo Mémoli, and Gunnar E Carlsson, "Topological methods for the analysis of high dimensional data sets and 3d object recognition.," in *SPBG*, 2007, pp. 91–100.

[12] M. Nicolau, A. Levine, and G. Carlsson, "Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival," *Proc. Natl. Acad. Sci.*, vol. 108, no. 17, pp. 7265–7270, 2011.

[13] P. Lum, G. Singh, A. Lehman, T. Ishkanov, M. Vejdemo-Johansson, M. Alagappan, J. Carlsson, and G. Carlsson, "Extracting insights from the shape of complex data using topology," *Sci. Rep.*, vol. 3, 2013.

[14] D. Hong, W. Liu, J. Su, Z. Pan, and G. Wang, "A novel hierarchical approach for multispectral palmprint recognition," *Neurocomputing*, vol. 151, pp. 511–521, 2015.

[15] A. Okujeni, S. Van Der Linden, and P. Hostert, "Berlin-urban-gradient dataset 2009-an enmap preparatory flight campaign (datasets)," *GFZ Data Services*, 2016.

[16] J. Hu, D. Hong, Y. Wang, and X. Zhu, "A comparative review of manifold learning techniques for hyperspectral and polarimetric sar image fusion," *Remote Sens.*, vol. 11, no. 6, pp. 681, 2019.

[17] D. Hong, N. Yokoya, and X. Zhu, "Learning a robust local manifold representation for hyperspectral dimensionality reduction," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 10, no. 6, pp. 2960–2975, 2017.