DOMAIN ADAPTATION WITH HIDDEN MARKOV RANDOM FIELDS

Jan-Pieter Jacobs[1*], *Guy Thoonen*[1], *Devis Tuia*[2], *Gustavo Camps-Valls*[3], *Birgen Haest*[4] *and Paul Scheunders*[1]

[1] iMinds-Vision Lab, University of Antwerp (Belgium)
[2] LaSIG laboratory, Ecole Polytechnique Fédérale de Lausanne (Switzerland)
[3] Image Processing Laboratory (IPL) Universitat de València (Spain)
[4] VITO-TAP, Remote Sensing Department, Flemish Institute for Technological Research (Belgium)

ABSTRACT

In this paper, we propose a method to match multitemporal sequences of hyperspectral images using Hidden Markov Random Fields. Based on the matching of the data manifold, the algorithm matches the reflectance spectra of the classes, thus allowing the reuse of labeled examples acquired on one image to classify the other. This allows valorization of spectra collected in situ to other acquisitions than the one they were acquired for, without user supervision, prior knowledge of the class reflectance in the new domain or global information about atmospheric conditions.

Index Terms— Multitemporal classification, domain adaptation, graph matching, Hidden Markov Random Fields.

1. INTRODUCTION

By the increased availability of hyperspectral (HS) imagery, many opportunities arise for detailed analysis of the surface of the Earth. One of the most common analysis methods to date is supervised classification, which allows detection of different types of land cover existing on the ground. A determining factor for reaching a satisfactory supervised classification result lies in the availability of sufficient ground reference data. Unfortunately, the collection of ground reference samples is often a time consuming and expensive task. As a result, many active and semi-supervised learning methods have been developed to aid in relaxing this requirement, either by selecting highly informative data points to be labeled by an expert [1], or by including characteristics of the unlabeled data in the classification process [2], respectively.

Independently from the acquisition method, available labeled samples are extremely valuable, and their reusability is of great interest: this is particularly true when considering images of the same area, at different moments in time, or when classifying geographically separate scenes that feature the same land cover classes. Reusing labeled samples is hindered by differences in illumination and atmospheric conditions between image acquisitions, that typically result in non-linear changes in the spectral signatures of equivalent materials. Consequently, in order to effectively reuse ground reference data, the images need to be radiometrically aligned, or *matched*, either in a physical way (through atmospheric correction) or by using relative and statistical methods founded in the field of machine learning called *Domain Adaptation (DA)* [3]. Since the former generally require additional ground and atmospheric data, which are often unavailable or hard to collect, the need for effective DA methods aligning image sequences for reusing labeled examples has become a reality in remote sensing (RS).

Until recently, most of the DA solutions in RS were based on integrating the adaptation directly in the classifier by using statistics of the new image. For instance, [4, 5] focus on the extension of kernel methods, such as the Support Vector Machine (SVM) classifier, with additional regularization terms, while [6] integrate the DA step into the active learning paradigm. By contrast, [7] introduces a generic unsupervised approach that directly transforms the underlying data structure, or *manifold* from one domain into another. To this end, a vector quantization algorithm extracts the local properties of the data clouds in both domains and two proximity graphs are built to represent the manifolds. Subsequently, these graphs are matched using a technique that minimizes the displacement of the graph's nodes, while preserving the local structure of the original graph. Afterwards, the training samples in the source domain are mapped to the target domain according to the match found before. Finally, the target scene is classified by an algorithm trained on the mapped training samples of the source.

We preserve the aforementioned vector quantization step and graph representation, but note that the approach of [7] is sensitive to larger manifold changes and variations in the graph representation. This issue stems from two causes: firstly, the use of Euclidean distances for assessing node similarity does not represent the manifold shape and secondly, the fact that the source and target graph's internal structures are not taken into account when matching the graphs. The former is remedied by using the geodesic distances over a k-Nearest Neighbor (k-NN) graph, in order to approximate the shape

*Email: jan-pieter.jacobs@ua.ac.be

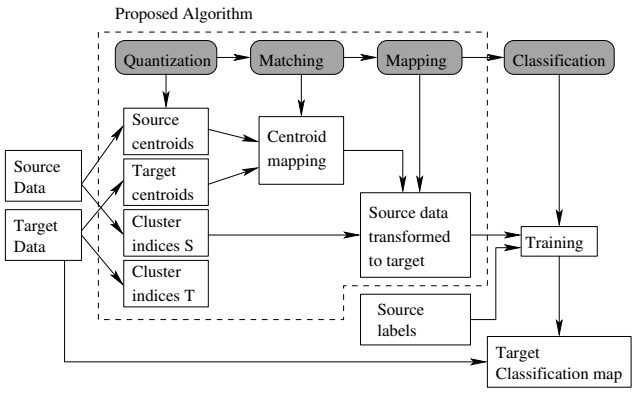


Fig. 1. Overview of the proposed algorithm.

of the data manifolds, while the latter is addressed by modeling the relationships between the nodes in feature space as a Hidden Markov Random Field (HMRF), considering the similarities among nodes both *across* graphs as *within* them. Because optimizing the problem as a full HMRF is NP-hard, we use a graphical model called Single Path Dynamic Programming (SPDP) [8], which globally optimizes the match between a sequence of target nodes and a chosen sequence of source nodes. After the matching, the cluster centroids in one domain are mapped upon their matches in the other domain (as in [7]), and classification is performed. We demonstrate the ability of the method to match multitemporal HS images by studying the correction of a large cloud shadow on the source image, reusing training data obtained for the source image to classify the target image.

2. METHODOLOGY

The proposed algorithm, illustrated in Fig. 1, consists of three main stages:

1. Vector Quantization by k-means clustering in each domain to obtain the centroids (Section 2.1)
2. Matching centroids between domains by SPDP (Section 2.2)
3. Mapping data according to matched centroids (Section 2.3)

2.1. Vector Quantization

The SPDP graph matching method, described in section 2.2, is a global optimization method. As such, the procedure is computationally intensive and scales badly with increasing image sizes. For this reason, we propose to use vector quantization prior to matching the domain manifold graphs, as is also done in [7]. To this end, the k-means clustering algorithm is applied to quantize local information into n cluster centroids; the subsequent matching procedure is only concerned with the centroids of the different domains.

2.2. Single Path Dynamic Programming

In the DA setting we consider source domain S (where we have labeled pixels) and target domain T (where we only use labeled pixels for testing), and represent them by the graphs $G_D = (\mathcal{V}_D, \mathcal{E}_D)$ with domain $D \in \{S, T\}$. $\mathcal{V}_D$ contains nodes x_D^i, i.e. reflectance spectra of centroids obtained by vector quantization. $\mathcal{E}_D$ contains edges w_D^{ij}, i.e. geodesic distances between nodes (x_D^i, x_D^j) over a k-NN graph from one point in feature space to another. By using geodesic distances the manifold shape is taken into account, improving the cross-domain representability [9].

We model the graphs for both domains as two observations of one common underlying HMRF. This model is described in terms of states and observations. States C_D are "hidden": it is not known in which state the system is, except for the observations x_D and w_D, which depend only on the states and transitions between them. These states, in turn, have an unknown influence on neighboring states, which is modeled through the observations. As such, the probability for any possible transition between states is assessed based on the observed outputs.

The HMRF probability model, as described by [8], is composed of 2 components:

The observation component $\phi_{i,k}$ is the probability that an observation x_S^i in the source has been generated by the same hidden state C_T^k that generated x_T^k in the target.

The Markov component $\psi_{ij,kl}$ is the probability that an observation of a binary attribute (edge) w_S^{ij} in the source has been generated by the same transition between hidden states C_T^k and C_T^l that generated w_T^{kl} in the target.

These probabilities, which indicate whether there is a match between nodes and edges of the graphs, are modeled by a similarity function. The similarity function used in this work, is the multivariate Gaussian probability density function, with x_T^i resp. w_T^{kl} as means and $\mathrm{cov}(x_T)$ resp. $\mathrm{cov}(w_T)$ as the covariance matrices, yielding:

$$\phi_{i,k} = p\left(X = x_S^i | X \sim \mathcal{N}\left(x_T^k, \mathrm{cov}\left(x_T\right)\right)\right) \quad (1)$$

$$\psi_{ij,kl} = p\left(W = w_S^{ij} | W \sim \mathcal{N}\left(w_T^{kl}, \mathrm{cov}\left(w_T\right)\right)\right) \quad (2)$$

Given these probabilities, the general problem of graph matching expressed in the HMRF model [8] is:

$$\begin{aligned} m^* &= \arg\max_m p(C_S^1 = C_T^{m(1)}, \ldots C_S^n = C_T^{m(n)} | G_S, G_T) \\ &= \prod_{i,j \in S} \phi_{i,m(i)} \cdot \psi_{ij,m(i)m(j)} \end{aligned} \quad (3)$$

where $m(i)$ is a function defining the mapping of source node x_S^i to target node $x_T^{m(i)}$ with optimum $m^*(\cdot)$ and n is the

number of source nodes $\#S$. $\phi_{i,m(i)}$ indicates the similarity of the mapped nodes and $\psi_{ij,m(i)m(j)}$ is the similarity of the edges between the mapped nodes.

In the SPDP algorithm by [8], the NP-hard problem posed above is reduced by constraining the HMRF to a single Markov chain, containing each source node exactly once. To this end, the algorithm considers a path through the source manifold, such that all nodes in the source domain are on the path exactly once. In this article, this path is the shortest path connecting all source nodes. To do so, we applied the Hungarian algorithm on the source graph's edges yielding a sequence of nodes $x_S^1, \ldots, x_S^n$ representing a path in the source graph. We aim to find the most probable corresponding sequence $C_T^{m(1)}, \ldots C_T^{m(n)}$, given the unary and binary compatibilities between observations and states (observation component) resp. between two states (Markov component). The algorithm is initialized by selecting the target cluster $C_T^{m(1)}$ matching the first source node in the path with the highest probability: $m(1) = \arg\max_k \phi_{1,k}$. For all following source nodes C_S^i in the path, we select a matching target state $C_T^{m(i)}$ maximizing

$$\begin{aligned} &\phi_{i,m(i)} \cdot \psi_{(i-1,i;m(i-1),m(i))} \\ &= p\left(x_S^i | C_T^{m(i)}\right) \cdot p\left(C_S^i = C_T^{m(i)} | C_S^{i-1} = C_T^{m(i-1)}\right) \quad (4) \end{aligned}$$

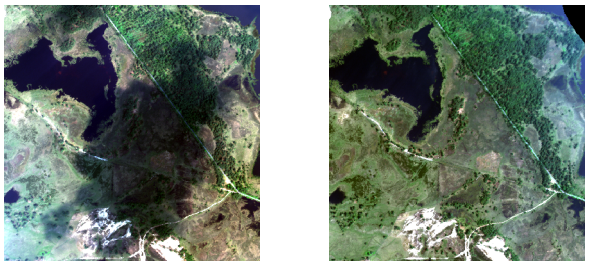
2.3. Mapping

Mapping data points from source to target is done by translating the data points in each cluster in the source by the amount of shift between the original cluster centroid and the mapped one, as in [7]. This adapts local information available in one cluster to match the data manifold in the target.

3. EXPERIMENTS AND RESULTS

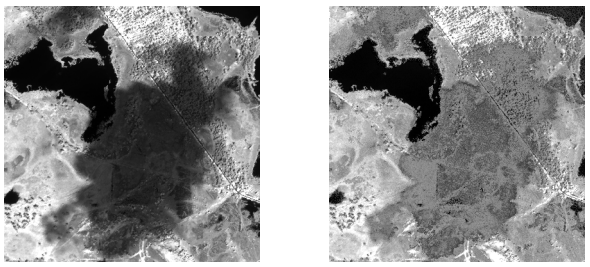
The proposed method was tested on HS imagery of a heathland area in Belgium, 'Kalmthoutse Heide'. For this area, airborne HS data were obtained in June 2007 with the Airborne Hyperspectral Scanner (AHS) sensor with a ground resolution of approximately 2.5m. The range of 450nm-2550nm is covered by 63 spectral bands, of which 11 noisy bands were removed prior to our experiments. Around 31000 ground reference points, over 5 classes, have been collected by photo-interpretation of a high resolution orthophotograph. About 1500 of these points are used for training. All classification results have been obtained using a SVM classifier with a RBF kernel. The two 400×400 pixels images under study, which serve as source and target, both describe the same scene, but are part of different flight lines. In the source, a large cloud shadow is visible (Fig. 2).

Fig. 3 shows band 12 of the source image before and after the mapping. Notice that, while the shadowed area is visibly adapted, other areas are not affected. Overall Accuracy (OA) and kappa values improve vastly when classifying the target image after training on the adapted source image. Table 1 compares the results after training on the original and mapped data, averaged over 10 random training sets.



(a) (b)

Fig. 2. True color representations of the AHS data, recording the same scene at a different time: (a) Source and (b) target.



(a) (b)

Fig. 3. Band 12 of the source image AHS data: (a) before and (b) after domain adaptation.

Table 1. Mean and standard deviation of OA, kappa and McNemar Z value vs. "No Action" over 10 test runs.

Mean (std.) of	OA (%)	kappa	McNemar Z
No Action	83.5 (3.30)	0.73 (0.051)	N/A
Method in [7]	88.1 (1.56)	0.80 (0.026)	19.1 (15.6)
Our Method	90.8 (1.06)	0.85 (0.017)	35.6 (13.2)

4. CONCLUSIONS

This article introduces a new procedure for domain adaptation. The approach consists of three main parts: A vector quantization step, a generic matching algorithm between the two domains, and a mapping step, in which the cluster centroids and all the data points contained within them, are mapped upon the matching cluster in the other domain. The graph matching algorithm uses a Hidden Markov Random Field model, and takes geodesic distances over a k-Nearest Neighbor graph accounting for the shape of the data manifold. Experiments on hyperspectral imagery demonstrate the effectiveness of our approach, which efficiently matches the data manifolds and corrects for a large shadow artifact. Visual improvements are backed up by classification improvements of about 7%.

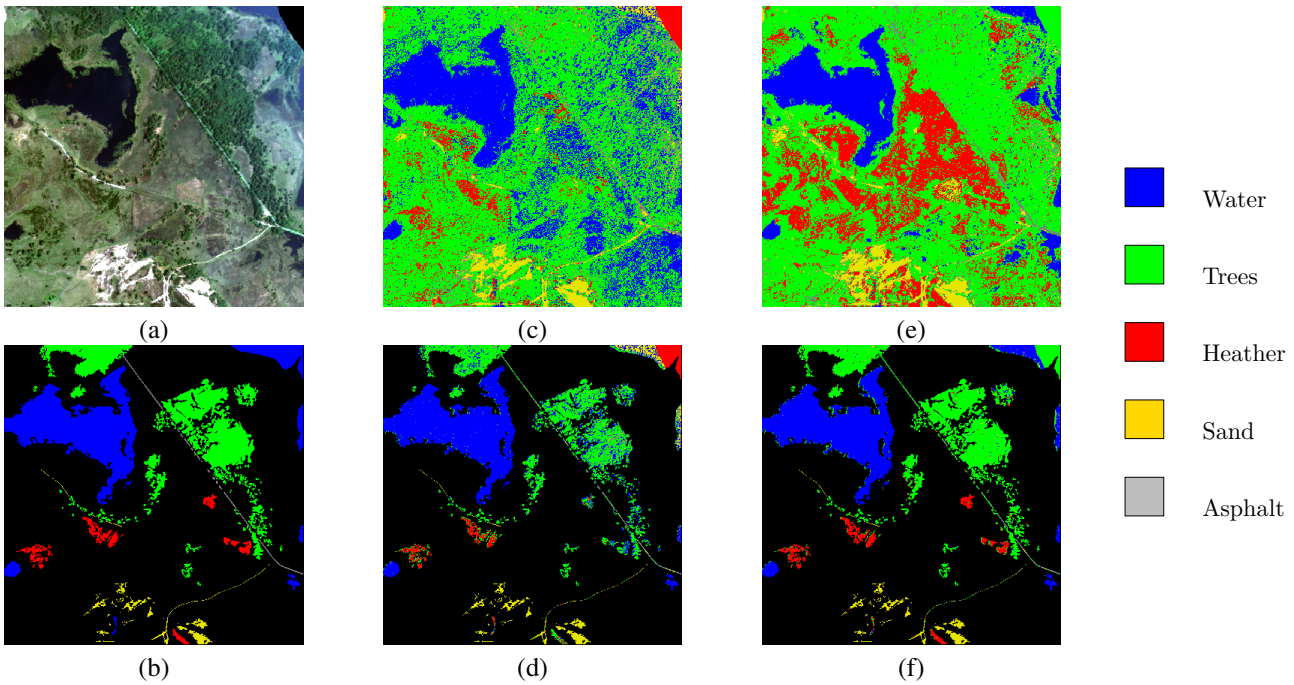


Fig. 4. Classification results. Left column: (a) true color representation of the target and (b) corresponding ground reference. Center column: (c) classification using original spectra on the entire image and (d) only on reference areas. Right column: (e) and (f) classification using the transformed spectra.

5. ACKNOWLEDGEMENTS

This work was funded by the Belgian Science Policy Office (Belspo), project RELEARN (contract no. SR/00/154), the Swiss National Foundation (grant PZ00P2-136827) and the Spanish Ministry of Economy and Competitiveness (MINECO project TIN2012-38102-C03-01). The authors would like to thank the Flemish Institute for Technological Research (VITO) for the acquisition and preprocessing of the AHS data.

6. REFERENCES

[1] D. Tuia, F. Ratle, F. Pacifici, M.F. Kanevski, and W.J. Emery, "Active learning methods for remote sensing image classification," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 47, no. 7, pp. 2218 –2232, july 2009.

[2] M. Seeger, "A taxonomy for semi-supervised learning methods," in *Semi-Supervised Learning*, O. Chapelle, B. Schölkopf, and A. Zien, Eds., pp. 15–31. MIT Press, Cambridge, MA, 2006.

[3] J. Quiñonero-Candela, M. Sugiyama, A. Schwaighofer, and N.D. Lawrence, *Dataset Shift in Machine Learning*, MIT Press, Cambridge, MA, 2009.

[4] L. Bruzzone and M. Marconcini, "Domain adaptation problems: A DASVM classification technique and a circular validation strategy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 5, pp. 770–787, 2010.

[5] W. Kim and M.M. Crawford, "Adaptive classification for hyperspectral image data using manifold regularization kernel machines," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 48, no. 11, pp. 4110 –4121, nov. 2010.

[6] G. Matasci, D. Tuia, and M. Kanevski, "SVM-based boosting of active learning strategies for efficient domain adaptation," *IEEE J. Sel. Topics Appl. Earth Observ.*, vol. 5, no. 5, pp. 1335–1343, 2012.

[7] D. Tuia, J. Muñoz Marí, L. Gómez-Chova, and J. Malo, "Graph matching for adaptation in remote sensing," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. PP, no. 99, pp. 1 –13, 2012.

[8] T. Caelli and T. Caetano, "Graphical models for graph matching: Approximate models and optimal algorithms," *Pattern Recognition Letters*, vol. 26, no. 3, pp. 339 – 346, 2005, In Memoriam: Azriel Rosenfeld.

[9] C.M. Bachmann, T.L. Ainsworth, and R.A. Fusina, "Exploiting manifold geometry in hyperspectral imagery," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 43, no. 3, pp. 441 –454, March 2005.