MULTICLASS DOMAIN ADAPTATION WITH ITERATIVE MANIFOLD ALIGNMENT

Brian D. Bue[1], *Chris Jermaine*[2]

Department of Electrical and Computer Engineering[1], Computer Science[2]
Rice University, Houston, TX, USA

ABSTRACT

We propose a novel approach for multiclass domain adaptation using an iterative manifold alignment technique inspired by the TRiplet-based Iterative ALignment (TRIAL) protein structure alignment algorithm. Our technique learns a rigid transformation for each class using a set of automatically-selected *pivot samples* that characterize the relative relationships between classes in two similar, but not identical, feature spaces. We demonstrate that our technique robustly reconciles domain-specific differences between similar classes in hyperspectral images captured under different conditions, and yields more accurate results than recently-proposed manifold alignment techniques. We evaluate our method on a pair of real-world hyperspectral images of Cuprite, NV, and provide a MATLAB implementation of our algorithm, available online.

Index Terms— domain adaptation, manifold alignment, multiclass, classification

1. DOMAIN ADAPTATION FOR HYPERSPECTRAL IMAGE CLASSIFICATION

Spectra captured by modern hyperspectral imaging platforms provide ample signal content to identify spectrally similar but distinct materials. Unfortunately, in remote-sensing applications, representative samples are not always available to train a classifier to reliably classify all known materials in a given scene. In such cases, incorporating training samples from other, similar images is an attractive option, but poses significant challenges. In particular, spectra representing identical materials have different spectral signatures when captured under different conditions (e.g., by different sensors, at different spatial locations, or at different capture times). Consequently, reconciling differences between training (or *source*) and test (*target*) spectra captured under different conditions is crucial to accurately transfer knowledge of source classes to classify target spectra.

Several recent works propose *domain adaptation* techniques to reconcile differences between spectra captured under different conditions. Some of these involve active learning techniques, which require user intervention to select target samples most relevant to the domain adaptation problem (e.g., [1, 2]). While active-learning approaches produce good results, requiring expert intervention limits the applicability of the technique for fully-autonomous applications, such as onboard spacecraft. Another approach is to automatically adapt a pre-trained classifier to classify similar imagery (e.g., [3, 4, 5]). However, such techniques assume a specific type of classifier has been trained that can subsequently be tuned to the new data.

An attractive alternative to the aforementioned techniques is to learn a transformation that maps the source and target spectra to a similar feature space. By learning such a mapping, we can apply a classification algorithm of our choosing in the transformed feature space. Yang et al. recently demonstrated that *manifold alignment* techniques are well-suited to learn such mappings for hyperspectral domain adaptation tasks [6]. However, most existing manifold alignment techniques learn a single *global* transformation between domains (e.g., [7, 8]. While applying a global transformation can resolve systematic differences between domains, it may prove inadequate in resolving the local differences caused by varying viewing geometries, illumination or atmospheric conditions that alter the radiances observed at the sensor of specific materials in a *class-specific* manner [9].

In this work, we present the MARTIAL (MAnifold Reconciliation Through Iterative ALignment) algorithm for multiclass domain adaptation. MARTIAL extends our previous domain adaptation work ([10]) by incorporating an iterative manifold alignment approach inspired by the TRIAL protein structure alignment algorithm of Venkateswaran et al. [11]. By learning a set of rigid transformations between the source and target feature spaces using a set of automatically-selected pivot samples representing identical classes in both domains, MARTIAL can reconcile class-specific differences more accurately than techniques that learn a single global transformation between domains. Additionally, we can train and apply any classifier in the resulting transformed feature space to classify target data. We evaluate our results on real-world hyperspectral images of Cuprite, NV, and provide a MATLAB implementation for convenient experimentation.[1]

2. MULTICLASS DOMAIN ADAPTATION WITH MARTIAL

We assume are provided N^S source domain samples $(\mathbf{X}^S, Y^S) = \left\{(\mathbf{x}_i^S, y_i^S)\right\}_{i=1}^{N^S}$, $\mathbf{x}_i^S \in \mathbb{R}^n$, $y_i^S \in \{1, \ldots, K\}$, drawn from source distribution $\mathrm{p}^S(\mathcal{X}, \mathcal{Y})$. We also assume $M^S \gg N^S$ unlabeled samples are available in the source domain $\mathbf{X}^{Su} = \left\{\mathbf{x}_i^{Su}\right\}_{i=1}^{M^S}$. Our goal is to find a transformation $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ that maps samples drawn from $\mathrm{p}^S(\mathcal{X}, \mathcal{Y})$ to the feature space of N^T target samples $\mathbf{X}^T = \left\{\mathbf{x}_i^T\right\}_{i=1}^{N^T}$, $\mathbf{x}_i^T \in \mathbb{R}^n$, drawn from a similar distribution $\mathrm{p}^T(\mathcal{X}, \mathcal{Y})$. We can subsequently train a classifier $h : \mathcal{X} \rightarrow \mathcal{Y}$ to predict the class labels Y^T using the transformed source samples $\mathbf{X}^{ST}$ as training data. The MARTIAL algorithm uses several components of the TRIAL algorithm to learn a transformation for each source class to the target domain. Before we describe the MARTIAL algorithm, we provide a brief synopsis of the TRIAL algorithm below.

The TRIAL Algorithm: Given a pair of proteins $\mathbf{A} = \{\mathbf{a}_i\}_{i=1}^{N^A}$ and $\mathbf{B} = \{\mathbf{b}_j\}_{j=1}^{N^B}$, each consisting of 3-dimensional C_α atoms $\{\mathbf{a}_i, \mathbf{b}_j\}$, TRIAL aligns the manifolds defined by $\mathbf{A}$ and $\mathbf{B}$ such that their *alignment length* – the number of paired C_α atoms $(\mathbf{a}_i, \mathbf{T} \cdot \mathbf{b}_j)$ nearby one another after applying a (3×3) transformation matrix $\mathbf{T}$ to each $\mathbf{b}_j \in \mathbf{B}$ – is maximized. This allows TRIAL to identify

[1] Available at: http://www.ece.rice.edu/~bdb1/#code

structural commonalities between $\mathbf{A}$ and $\mathbf{B}$ which may be arbitrarily rotated with respect to one another. The algorithm consists of three main steps: (1) triplet (seed) alignment (denoted *Seed*), (2) initial alignment (*Align*), and (3) iterative improvement (*Improve*). In step (1), TRIAL searches for pairs of *triplet* (or *seed*) C_α atoms $\mathbf{P}^A = \{\mathbf{p}_1^A, \mathbf{p}_2^A, \mathbf{p}_3^A\} \subset \mathbf{A}$, $\mathbf{P}^B = \{\mathbf{p}_1^B, \mathbf{p}_2^B, \mathbf{p}_3^B\} \subset \mathbf{B}$ that are structurally similar to one another in terms of the Euclidean distances between their constituent atoms. TRIAL uses these pairs to find a preliminary, minimum root mean square deviation (RMSD) alignment between $\mathbf{A}$ and $\mathbf{B}$ using the Kabsch algorithm [12]. After filtering out any seed pairs producing degenerate (i.e., high RMSD) alignments, TRIAL reduces the RMSD between $\mathbf{P}^A$ and $\mathbf{P}^B$ while increasing the number of atoms in alignment (step (2)). It achieves this by iteratively recomputing $\mathbf{T}$ after adding any pairs $(\mathbf{a}_i, \mathbf{T} \cdot \mathbf{b}_j)$ whose Euclidean distances are less than a user-defined threshold ϵ to the $\mathbf{P}^A$ and $\mathbf{P}^B$ sets, repeating the process until no more such pairs within the threshold can be added. During the final, iterative improvement step (3), TRIAL ensures that alignment length between $\mathbf{A}$ and $\mathbf{B}$ is maximized without increasing the RMSD of the aligned solution. Similarly to step (2), this involves recomputing $\mathbf{T}$ after iteratively adding any $(\mathbf{a}_i, \mathbf{T} \cdot \mathbf{b}_j)$ with distance less than an upper bound ϵ_{max} computed from the current solution set $(\mathbf{P}^A, \mathbf{P}^B, \mathbf{T})$.

Domain Adaptation with the MARTIAL Algorithm: The TRIAL algorithm has several attractive properties that lend themselves favorably to domain adaptation problems. Whereas several existing manifold alignment techniques assume a substantial quantity of (labeled) pairwise correspondences between domains are available at initialization (e.g. [7, 8]), TRIAL is capable of adapting to the properties of the source and target manifolds with a relatively small number of labeled correspondences ($\approx 10-100$ per-class) by iteratively refining the mapping between domains by incorporating informative *unlabeled* samples. Additionally, the rigid transformations computed by TRIAL preserve functional relationships between adjacent spectral bands, which are crucial for accurate classification of hyperspectral signatures [13].

However, several issues arise which prevent us from applying TRIAL directly in domain adaptation scenarios. Specifically, in domain adaptation, our objective is to minimize misclassifications, rather than maximizing the number of aligned samples between the source and target domains. Additionally, while we can assume that the C_α atoms in $\mathbf{A}$ and $\mathbf{B}$ each lie on single submanifold of $\mathbb{R}^3$, samples representing different classes in the source and target data can be viewed as lying on their own submanifolds of $\mathbb{R}^n$, and the submanifold of a particular class in the target domain may be arbitrarily transformed with respect to the submanifold of the same class in the source domain. Finally, we must consider problems involving hundreds to thousands of samples of high dimensionality, which involves significantly greater computational costs than those involved in protein alignment problems.

We account for these challenges by making the following modifications the TRIAL algorithm: (1) we perform an initial filtering step where we select a pool of candidate pivot samples that are structurally similar with respect to class structure in both the source and target domains; (2) rather than learning a single global transformation between the domains, we learn separate transformations for each of the source classes using the candidate pivot samples. This allows us to resolve domain-specific differences relative to each class, while also constraining the number of samples considered during alignment; and (3) we automatically compute the RMSD threshold ϵ for each class by randomly selecting a set of initial *seed* pairs of fixed size from the set of candidate pivots that produce a low RMSD transform between the pivots for each class. While this is not guaranteed to produce an optimal RMSD transformation, we found that selecting the lowest RMSD transformation over 25-50 seed pairs works well in practice to filter out degenerate solutions.

Algorithm 1 MARTIAL

Input: N^S labeled source samples $(\mathbf{X}^S, Y^S)$, M^S unlabeled source samples $\mathbf{X}^{Su}$, N^T unlabeled target samples $\mathbf{X}^T$, number of candidate pivots N_i^P per class, number of seed samples per class Q_i, number of random inits N_{rand}.

Output: Target-transformed source samples $\mathbf{X}^{ST}$

1: Use MCCL to select N^P candidate pivots $\mathbf{P} = (\mathbf{P}^S, \mathbf{P}^T, Y^P)$, $\mathbf{P}^S \subset (\mathbf{X}^S \cup \mathbf{X}^{Su})$, $\mathbf{P}^T \subset \mathbf{X}^T$.
2: $\mathbf{X}^{ST} = \emptyset$
3: **for** $i = 1$ **to** K **do**
4: $\quad \mathbf{X}_i^S = \{\mathbf{x}_j^S \in \mathbf{X}^S : y_j^S = i\}$
5: $\quad \mathbf{P}_i = \{(\mathbf{p}_j^S, \mathbf{p}_j^T) \in \mathbf{P} : y_j^P = i\}$
6: $\quad (\mathbf{P}_{\text{Seed}}, \mathbf{T}_{\text{Seed}}, \epsilon_{\text{Seed}}) = \text{SEEDINIT}(\mathbf{P}_i, Q_i, N_{\text{rand}})$
7: $\quad \mathbf{T}_i^S = \text{TRIAL}(\mathbf{P}_i, \mathbf{P}_{\text{Seed}}, \mathbf{T}_{\text{Seed}}, \epsilon_{\text{Seed}})$
8: $\quad \mathbf{X}^{ST} = \{\mathbf{X}^{ST} \cup \mathbf{T}_i^S \cdot \mathbf{X}_i^S\}$
9: **end for**

Algorithm 1 describes the MARTIAL algorithm, which maps a set of labeled source samples $(\mathbf{X}^S, Y^S)$ to the target domain feature space. The algorithm begins by using the Multiclass Continuous Correspondence Learning (MCCL) algorithm ([10], Algorithm 1) to select a pool $\mathbf{P} = (\mathbf{P}^S, \mathbf{P}^T, Y^P)$ of N^P candidate pivot samples. We denote the set of N_i^P pivots representing the i^{th} class as $\mathbf{P}_i = (\mathbf{P}_i^S, \mathbf{P}_i^T) = \{(\mathbf{p}_j^S, \mathbf{p}_j^T)\}_{j=1}^{N_i^P}$. The set of N^P source pivots $\mathbf{p}_j^S \in \mathbf{P}^S$ consist of the top N_i^P samples in $(\mathbf{X}^S \cup \mathbf{X}^{Su})$ nearest to the mean of each source class. For each source pivot $\mathbf{p}_j^S \in \mathbf{P}^S$, MCCL selects the target pivot $\mathbf{p}_j^T = \mathbf{x}_\ell^T \in \mathbf{X}^T$ most likely to belong to the same class as $\mathbf{p}_j^S$ according to

$$\ell = \underset{i}{\text{argmin}} \, \|\text{R}(\mathbf{p}_j^S, \mathbf{P}^S) - \text{R}(\mathbf{x}_i^T, \mathbf{P}^S)\|, \; i \in \{1, \ldots, N^T\}, \quad (1)$$

where

$$\text{R}(\mathbf{x}, \mathbf{P}^D) = \left(\frac{\text{d}(\mathbf{x}, \mathbf{p}_1^D)}{\sum_{\ell=1}^Q \text{d}(\mathbf{x}, \mathbf{p}_\ell^D)}, \ldots, \frac{\text{d}(\mathbf{x}, \mathbf{p}_Q^D)}{\sum_{\ell=1}^Q \text{d}(\mathbf{x}, \mathbf{p}_\ell^D)} \right), \quad (2)$$

for $D \in \{S, T\}$ and $Q = \sum_{i=1}^K Q_i$. The R function maps $\mathbf{x}$ to a Q-dimensional feature space where each feature gives the likelihood of distinguishing $\mathbf{x}$ from pivot sample $\mathbf{p}_\ell^D \in \mathbf{P}^D$, according to distance measure $\text{d}(\cdot, \cdot)$[2]. By selecting the candidate pivots in this "R-space," MCCL finds target samples that approximately preserve the relative distances between the source classes. When the source and target feature spaces are similar, these target pivots typically represent the same classes as their corresponding source pivots.

After selecting the candidate pivots, MARTIAL uses $\mathbf{P}_i$ to compute the Seed alignment transformation $\mathbf{T}_{\text{Seed}}$ for source samples from the i^{th} class, X_i^S. This is achieved by sampling N_{rand} seed pairs from $\mathbf{P}_i$, each consisting of $Q_i < N_i^P$ samples of the form $\mathbf{P}_{\text{Seed}} = (\mathbf{P}_{\text{Seed}}^S, \mathbf{P}_{\text{Seed}}^T) = \{(\mathbf{p}_j^S, \mathbf{p}_j^T)\}_{j=1}^{Q_i}$, and applying the Kabsch algorithm to each seed pair, returning the $(\mathbf{T}_{\text{Seed}}, \mathbf{P}_{\text{Seed}})$ that yields the smallest value of $\epsilon_{\text{Seed}} = \text{RMSD}(\mathbf{T}_{\text{Seed}} \cdot \mathbf{P}_{\text{Seed}}^S, \mathbf{P}_{\text{Seed}}^T)$ (Step 6).

We then pass this filtered set of pivots to the TRIAL function for refinement (Step 7), which performs the initial alignment and iterative improvement steps of the TRIAL algorithm as described in [11] (Figures 2 and 4). The TRIAL function returns the $n \times n$ transformation matrix $\mathbf{T}_i^S$ that maps $\mathbf{X}_i^S$ target feature space. We

[2] In this work, $\text{d}(\cdot, \cdot)$ is the Euclidean distance.

add $\mathbf{T}_i^S \cdot \mathbf{X}_i^S$ to the set of transformed source samples $\mathbf{X}^{ST}$ (Step 8). We can then use the transformed source samples $(\mathbf{X}^{ST}, Y^S)$ to train a multiclass classifier to predict labels for target samples $\mathbf{X}^T$.

3. EXPERIMENTAL RESULTS: CUPRITE, NV IMAGERY

We consider a challenging multi-sensor, multi-temporal domain adaptation problem to evaluate our methodology. Our objective is to classify a set of mineralogical spectra from one hyperspectral image using training data from another image of the same location, but captured by a different sensor and under different environmental conditions. Our data consists of five mineralogical classes manually labeled by an expert geologist from two images of the Cuprite mining district in Cuprite, NV. Image *Av97* was captured in June 19, 1997 by the AVIRIS instrument, consists of 512×614 pixels, and was studied in detail in [14]. Image *Hyp11* was acquired on Feb. 6, 2011 by the Hyperion instrument onboard the EO-1 satellite, and contains 1798×779 pixels. Each pixel is a 29-dimensional vector of image radiance values measured at wavelengths in the range 2.1029-2.3249μm. We perform atmospheric calibration via the empirical line method and scale each pixel by its L^2 norm to compensate for linear illumination effects. The class means and sample counts from each image are shown in Figure 1. We segment each image using the technique described in [15], and use the means of the resulting segments for pivot selection. Initial domain adaptation experiments, along with false-color composites of each image and training sample locations are described in detail in [10].

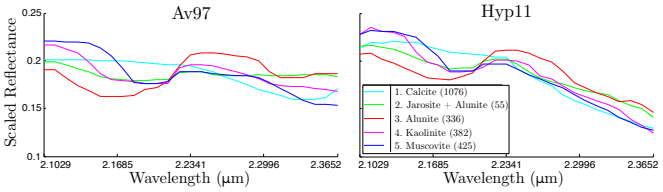

Fig. 1. Class means for Av97 (left) and Hyp11 (right) images. Sample counts for each class are given in parenthesis.

We consider the following two scenarios. We first train a classifier using source data from the Av97 image to classify target data from the Hyp11 image. We refer to this scenario as *Av97⇒Hyp11* . In the second scenario, *Hyp11⇒Av97* , we use the Hyp11 data as the source data, and classify target data from the Av97 image. In each scenario, we measure the baseline (i.e., no domain adaptation) source-to-target (ST) classification accuracy, which provides a baseline accuracy we seek to improve. We then measure the classification accuracy using the transformations produced by the MARTIAL Seed (Algorithm 1, Step 6), Align and Improve (Step 7) steps. We select $N_i^P = 250$ candidate pivots from each class, and evaluate classification accuracy for seed sizes $Q_i \in \{10, 12, 15, 20, 24, 30, 36, 40, 42, 50, 75, 100\}$. We compare our results to those produced using the "Manifold Alignment using Procrustes Analysis" (denoted *Procrustes*) technique of Wang and Mahadevan [8], which computes a single *global* transformation between the source and target domains using the Kabsch algorithm.[3] We also provide results after mapping the source and target spectra to the R-space using source samples in their original feature space (R_{Source}) and the source samples produced after applying the MARTIAL Seed (R_{Seed}), Align (R_{Align}) and Improve ($R_{Improve}$) steps. We use the same Q_i pivots from each class used in the MARTIAL Seed alignment step for the Procrustes and the R-space mappings. Our classifier is the multiclass linear Support Vector Machine (SVM) implemented in the LIBSVM package [16], evaluated using five-fold cross-validation. We select the SVM slack parameter $C \in \{10^{-3}, \ldots, 10^3\}$ that yields the highest accuracy on the source data.

Figure 2 shows the classification accuracy vs. the number of seed samples Q_i for each algorithm in the Av97⇒Hyp11 (left) and Hyp11⇒Av97 (right) scenarios. In the Av97⇒Hyp11 scenario, we observe that classifying source samples after each of the MARTIAL Seed, Align and Improve steps produces accuracies significantly better than the baseline (8-11%). The poor performance by the Procrustes alignment algorithm for most Q_i values suggests that the single global transformation computed using the pivot samples does not adequately resolve the class-specific differences between the images. We also observe dramatic improvements over the Procrustes alignment using MARTIAL in the Hyp11⇒Av97 scenario. However, as noted in [10], because the classes are better separated in the Av97 image than in the Hyp11 image, we achieve high classification accuracy ($\approx 94\%$) in the Hyp11⇒Av97 scenario with the baseline (ST) classifier. The remaining classes are challenging to separate, as indicated by the roughly comparable performance to the baseline using each of the domain adaptation algorithms. On average, however (as shown in Table 1 below), classifying source samples transformed by MARTIAL yields slightly better accuracies than the baseline.

We observe more substantial improvements in classification accuracy when we classify our data in the R-space (Equation (2)) after applying MARTIAL (Figure 3). In the Av97⇒Hyp11 scenario, classifying the target samples in the R-space using the source data transformed by MARTIAL produces uniformly better results for all Q_i than in the R-space with the original source features (R_{source}), indicating that the domains are better reconciled after applying the MARTIAL transformations. The R-space classification results using MARTIAL are also better than those given in Figure 2 for all $Q_i \neq 100$. Not surprisingly, as the classification accuracies in the Hyp11⇒Av97 scenario are already high, the R_{Source} and the MARTIAL R_{Align} and $R_{Improve}$ cases produce comparable, but not significantly better accuracies ($\pm 1\%$).

Table 1 provides a summary of the classification accuracies of each method, averaged over the range of Q_i values. We see that the MARTIAL feature space produced by the Align step yield the most accurate results in the Av97⇒Hyp11 scenario, and perform comparably to MCCL in the Hyp11⇒Av97 scenario. We also note that the accuracies produced after applying the Align step are typically equal or slightly better than those produced after the subsequent Improve step. This may be somewhat surprising, as one may expect that incorporating additional samples in the Improve step would produce a more robust alignment between the domains. However, since the pivots from each class are highly-correlated, using a large number of redundant pivots often produces worse results than using a smaller set of less-redundant pivots.

4. CONCLUSIONS AND FUTURE WORK

In this work, we introduced the MARTIAL algorithm, which extended the TRIAL protein alignment algorithm to high-dimensional, multiclass domain adaptation scenarios. We showed improvements in classification accuracy of up to 10% over manifold alignment techniques that learn a single global transformation between domains by learning a set of transformations for each class according to a set of automatically selected pivot samples, and up to 6% improvements over our previously-proposed domain adaptation technique, MCCL. Our experiments indicate our technique typically yields accuracies slightly above (0.5-1.5%) the baseline (ST) in

[3] The MARTIAL seed alignment step can be interpreted as applying the Procrustes alignment algorithm to the pivots representing each class.

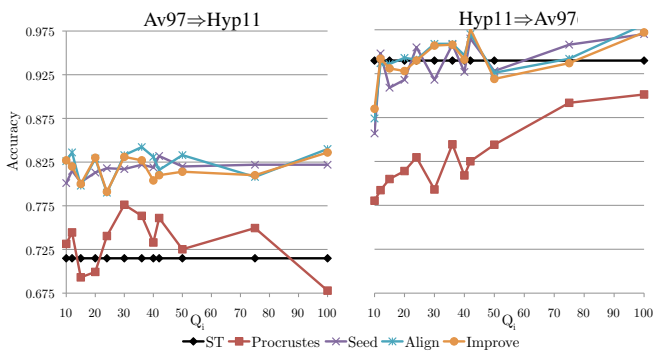

Fig. 2. Classification accuracy vs. number of seed samples Q_i for the Av97⇒Hyp11 (left) and Hyp11⇒Av97 (right) scenarios with the baseline (ST, black ◇), Procrustes alignment (red □), and MARTIAL Seed (purple ×), Align (turquoise ∗), and Improve (orange ○).

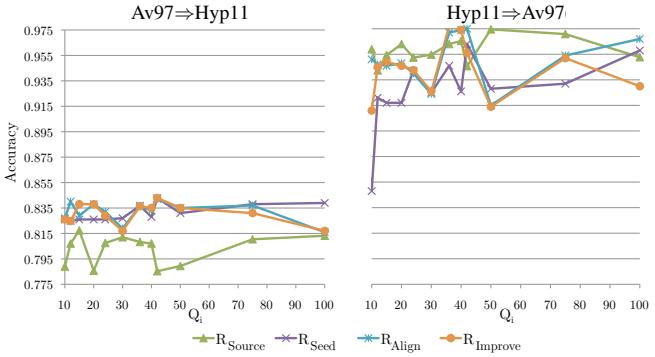

Fig. 3. R-space classification accuracy vs. number of seed samples Q_i for the Av97⇒Hyp11 (left) and Hyp11⇒Av97 (right) scenarios using source samples from the original source feature space (R_{Source}, green △) vs. the MARTIAL seed (R_{Seed}, purple ×), align (R_{Align}, turquoise ∗) and improve ($R_{Improve}$, orange ○) feature spaces.

	ST	**Procrustes**	**Seed**	**Align**	**Improve**	**R_{Source}**	**R_{Seed}**	**R_{Align}**	**$R_{Improve}$**
Av97⇒Hyp11	71.49%	73.29%	81.69%	82.35%	81.67%	80.27%	*83.10%*	*83.20%*	83.08%
Hyp11⇒Av97	93.99%	82.75%	93.43%	94.28%	94.05%	*95.82%*	92.69%	*95.11%*	94.36%

Table 1. Average accuracy over the range of selected Q_i values for each technique. The first and second most accurate results are given in red and blue italics, respectively.

cases when the baseline accuracy is already high.

While our algorithm performs well for a range of seed sizes Q_i, a question remains on how to automatically select a good value of Q_i for a particular domain adaptation problem. Initial experiments indicate that, in most cases, we can select an acceptable value Q_i for the task using a technique similar to the Pdiv model selection technique described in [10]. A forthcoming publication will detail these results.

Acknowledgement: The authors thank Chang Wang for his assistance with the Procrustes manifold alignment technique in the early stages of this work.

5. REFERENCES

[1] M. M. Crawford, D. Tuia, and H. L. Yang, "Active Learning: Any Value for Classification of Remotely Sensed Data?" *Proceedings of the IEEE*, vol. 101, no. 3, pp. 593–608, Mar. 2013.

[2] C. Persello and L. Bruzzone, "A novel active learning strategy for domain adaptation in the classification of remote sensing images," *IEEE Geoscience and Remote Sensing Symposium*, pp. 3720–3723, 2011.

[3] S. Rajan, J. Ghosh, and M. M. Crawford, "Exploiting Class Hierarchies for Knowledge Transfer in Hyperspectral Data," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3408–3417, 2006.

[4] W. Kim and M. M. Crawford, "Adaptive classification for hyperspectral image data using manifold regularization kernel machines," *Geoscience and Remote Sensing*, 2010.

[5] L. Bruzzone and M. Marconcini, "Domain Adaptation Problems: A DASVM Classification Technique and a Circular Validation Strategy," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 770–787, 2010.

[6] H. L. Yang and M. M. Crawford, "Manifold Alignment For Classification Of Multitemporal Hyperspectral Data," *Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, pp. 1–4, Apr. 2011.

[7] J. Ham, D. Lee, and L. K. Saul, "Semisupervised alignment of manifolds," *Proceedings of the Annual Conference on Uncertainty in Artificial Intelligence, Z. Ghahramani and R. Cowell, Eds*, vol. 10, pp. 120–127, 2005.

[8] C. Wang and S. Mahadevan, "Manifold alignment using Procrustes analysis," *Proceedings of the 25th international conference on Machine learning*, pp. 1120–1127, 2008.

[9] J. Adams and A. Gillespie, *Remote sensing of landscapes with spectral images: A physical modeling approach*. Cambridge Univ Pr, 2006.

[10] B. D. Bue and D. R. Thompson, "Multiclass Continuous Correspondence Learning," *NIPS Domain Adaptation Workshop*, Dec. 2011.

[11] J. Venkateswaran, B. Song, T. Kahveci, and C. Jermaine, "TRIAL: A Tool for Finding Distant Structural Similarities," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 8, no. 3, pp. 819–831, 2011.

[12] W. Kabsch, "A discussion of the solution for the best rotation to relate two sets of vectors," *Acta Crystallographica Section A: Crystal Physics*, vol. 34, pp. 827–828, Sep. 1978.

[13] T. Villmann, E. Merényi, and B. Hammer, "Neural maps in remote sensing image analysis," *Neural Networks*, vol. 16, no. 3-4, pp. 389–403, 2003.

[14] F. A. Kruse, J. Boardman, and J. Huntington, "Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1388–1400, 2003.

[15] D. R. Thompson, L. Mandrake, M. S. Gilmore, and R. Castaño, "Superpixel endmember detection," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–19, Jun 2010.

[16] C. Chang and C. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, no. 3, p. 27, 2011.